

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:29:52 ; Search time 18 Seconds

(without alignments)
711.054 Million cell updates/sec

Title: US-09-913-414-5

Perfect score: 2165

Sequence: 1 MKPPKIQRPDMFOAVARIG.....VLQYLFSPRLRTFTREE 435

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued Patents-AA:*

2: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*

3: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*

4: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*

5: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*

6: /cgn2_6/ptodata/1/1aa/6C.COMB.pep:*

7: /cgn2_6/ptodata/1/1aa/6D.COMB.pep:*

8: /cgn2_6/ptodata/1/1aa/6E.COMB.pep:*

9: /cgn2_6/ptodata/1/1aa/6F.COMB.pep:*

10: /cgn2_6/ptodata/1/1aa/6G.COMB.pep:*

11: /cgn2_6/ptodata/1/1aa/6H.COMB.pep:*

12: /cgn2_6/ptodata/1/1aa/6I.COMB.pep:*

13: /cgn2_6/ptodata/1/1aa/6J.COMB.pep:*

14: /cgn2_6/ptodata/1/1aa/6K.COMB.pep:*

15: /cgn2_6/ptodata/1/1aa/6L.COMB.pep:*

16: /cgn2_6/ptodata/1/1aa/6M.COMB.pep:*

17: /cgn2_6/ptodata/1/1aa/6N.COMB.pep:*

18: /cgn2_6/ptodata/1/1aa/6O.COMB.pep:*

19: /cgn2_6/ptodata/1/1aa/6P.COMB.pep:*

20: /cgn2_6/ptodata/1/1aa/6Q.COMB.pep:*

21: /cgn2_6/ptodata/1/1aa/6R.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	521.5	24.1	443	2	US-09-005-232A-3
2	510.5	23.6	443	2	US-08-620-605D-3
3	313.5	14.5	477	3	US-08-772-270A-13
4	262	12.1	478	3	US-08-772-270A-5
5	201.5	9.3	425	2	US-08-853-659A-38
6	118.5	5.5	796	4	US-08-868-699A-2
7	118.5	5.5	796	4	US-09-757-014-2
8	115	5.3	1090	4	US-09-085-199B-5
9	113.5	5.2	457	4	US-08-924-629C-6
10	112	5.2	531	4	US-09-134-001C-4920
11	109	5.0	1664	1	US-09-085-199B-4
12	109	5.0	1664	1	US-09-599-652-2
13	109	5.0	1664	2	US-08-642-846-2
14	109	5.0	1664	2	US-09-264-604-2
15	107	4.9	747	2	US-08-816-693A-51
16	107	4.9	747	2	US-08-885-291-51
17	107	4.9	747	4	US-09-496-672-51
18	106.5	4.9	591	1	US-08-178-738-5
19	106.5	4.9	591	2	US-08-628-145-5
20	106	4.9	2285	2	US-09-308-375-2
21	105	4.8	693	4	US-08-235-836C-68
22	104	4.8	728	4	US-09-134-001C-4968
23	104	4.8	1146	4	US-08-914-999-6
24	103	4.8	379	1	US-08-279-270A-1
25	102.5	4.7	529	4	US-08-887-534A-74
26	102.5	4.7	635	4	US-09-341-833A-7
27	102.5	4.7	677	4	US-09-341-833A-8

28	102.5	4.7	1048	4	US-08-887-534A-85	Sequence 85, Appl
29	102.5	4.7	1780	1	US-08-769-309A-5	Sequence 5, Appl
30	102.5	4.7	1780	3	US-08-994-570-5	Sequence 5, Appl
31	102	4.7	659	4	US-08-894-818B-1	Sequence 1, Appl
32	102	4.7	659	4	US-09-445-472-12	Sequence 12, Appl
33	102	4.7	756	4	US-09-085-199B-9	Sequence 9, Appl
34	102	4.7	896	1	US-08-095-737-2	Sequence 2, Appl
35	102	4.7	896	1	US-08-480-145-2	Sequence 2, Appl
36	102	4.7	896	2	US-08-477-389-2	Sequence 2, Appl
37	101.5	4.7	534	4	US-09-103-664A-2	Sequence 2, Appl
38	101.5	4.7	693	4	US-08-235-836C-72	Sequence 72, Appl
39	101	4.7	635	4	US-09-341-833A-9	Sequence 4, Appl
40	101	4.7	1388	2	US-08-685-576-4	Sequence 4, Appl
41	101	4.7	2101	1	US-08-195-487-4	Sequence 4, Appl
42	101	4.7	2101	5	PCT-US93-06160-4	Sequence 4, Appl
43	100.5	4.6	635	4	US-09-341-833A-6	Sequence 6, Appl
44	100	4.6	1388	2	US-08-685-576-1	Sequence 1, Appl
45	99.5	4.6	955	1	US-08-006-676B-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-005-232A-3
Sequence 3, Application US/09005232A
Patent No. 5922568
GENERAL INFORMATION:
APPLICANT: SHIBATANI, TAKEJI
APPLICANT: AKATSUKA, HIROYUKI
APPLICANT: KAWAI, ERI
TITLE OF INVENTION: GENE PARTICIPATING IN THE MECHANISM OF
TITLE OF INVENTION: SECRETION OF ESTERASE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESS: BIRCH, STEWART, KOLASCH AND BIRCH
STREET: PO BOX 747
CITY: FALLS CHURCH
STATE: VA
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/005, 232A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: SVENSSON, LEONARD R
REGISTRATION NUMBER: 30,330
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)-205-8000
TELEFAX: (703)-205-8050
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 443 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-005-232A-3
Query Match
Best Local Similarity 29.3%, Pred. No. 4.1e-41;
Matches 133; Conservative 88; Mismatches 194; Indels 39; Gaps 7;
OY 6 IORPTDNF-----QAVARIGYIATLVYGLGMAFAPLDSAVTANGVSAEVS-Q 56
DB 5 IGEPODSTTEIPQDERFRTRGMVLVYIGLGFPLAAMAFAPLQKGVASPGSVTVSCNRK 64

```

; SEQUENCE CHARACTERISTIC
; LENGTH: 443 amino acid
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-620-605D-3

```

REGISTRATION NUMBER: 40,261
REFERENCE/DOCKET NUMBER: 6580-81
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398
INFORMATION FOR SEQ ID NO: 13:

ADDRESS: Intellectual Property Services
ADDRESS: Battelle Memorial Institute
ADDRESS: PNL P.O. Box 999
STREET: Washington Way


```

STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103-2793
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/757.014
FILING DATE: 09-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/868.699
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Falk, Stephen T
REGISTRATION NUMBER: 36,795
REFERENCE/DOCKET NUMBER: GM10012
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-994-2488
TELEFAX: 215-994-2222
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 796 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-757-014-2

Query Match      5.5%; Score 118.5; DB 4; Length 796;
Best Local Similarity 20.6%; Pred. No. 0.013;
Matches 63; Conservative 52; Mismatches 120; Indels 71; Gaps 11;

QY 10 TDNFOAVARIGGIALFVGLGMAAFPLDSAVIANGVSAEVSODVQH----- 61
DB 358 TITFONLKRLEFSFGMTATGLGSEFFDLYSKI-----VVOAPFDKAIQRIDEDPKYFR 413
QY 62 -----EGMLAKILVREGKVK-AGOVLELDEPTQANAAAGITRNOYV 103
DB 414 SVDEKNIAMHDIHELHETGRPVLLITRTAEAEYFEVLQMD-----IPNNLLI 464
QY 104 ALK-AMEARLLAEQRQPSISFPADLTQSQRADPMVARRAIADQAFTERQTIQOVULM 162
DB 465 AONVAKEAQMIAEAGQISMTVATSMAGRTDIKLGEGV-----EALGGLAVII 513
QY 163 NAORLOVSEIEGIDRQGLKDOLG-----FIEDELIDRLKLYDKGLVPRPRLIA 213
DB 514 HE-----HMENSRVDRQLRGSRGQDPGSSCIYISLDDYLV--KMSDSNLAENNOQLYS 566
QY 214 LEARAGLSIGSIGRLTADRASKAVGASDTQ-LKVRQIKOEFEQVYS---QSITETRVRLA 269
DB 567 LDAQRLSQSNLFRKVKQIVVAKQRISEBGYKAREMANEFKKSISIQDLYEERNRVL 626
QY 270 EYTEKE 275
DB 627 EIDDAE 632

```

RESULT 8
US-09-085-199B-5
Sequence 5, Application US/09085199B
Patent No. 6235879

GENERAL INFORMATION:
APPLICANT: Hayden, Michael R.
APPLICANT: Hayden, Abigail
APPLICANT: Hug, A.H.M. Mahabul
APPLICANT: Chopra, Vikramjit Singh
APPLICANT: Kalchman, Michael

```

TITLE OF INVENTION: Apoptosis Modulators That Interact with the
TITLE OF INVENTION: Huntington's Disease Gene
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESS: Opedahl & Larson
STREET: PO Box 5270
CITY: Frisco
STATE: CO
COUNTRY: USA
ZIP: 80443-5270
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Kb storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS DOS 5.0
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/085.199B
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Larson, Marina T.
REGISTRATION NUMBER: 32038
REFERENCE/DOCKET NUMBER: UBC.P-013052
TELECOMMUNICATION INFORMATION:
TELEPHONE: (970) 668-2050
TELEFAX: (970) 668-2052
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1090
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: no
ORIGINAL SOURCE:
ORGANISM: human
FEATURE:
OTHER INFORMATION: Huntington-Interacting protein
US-09-085-199B-5

```

Query Match 5.3%; Score 115; DB 4; Length 1090;
Best Local Similarity 20.1%; Pred. No. 0.045;
Matches 82; Conservative 72; Mismatches 131; Indels 122; Gaps 20;

```

QY 47 NGVSAEVSODVQHE---GMLAKILVREGKVKAGOVLELDEPTQANAAAGITRNOYV 103
DB 419 NGVNKDEKDHILRLRLREISGLKQL---ENKKTESQRYVLQKGVSELEADLDAQHL 475
QY 104 ALKAME--ARLIAERDQRPISFPADLTQSQRADPMVARRAIADQAFTERR----- 152
DB 476 RQQAADDEFLRAELD-----ELRRQREDTEKQARSISEIRKQANQORYSKILKE 526
QY 153 ---QTIQOVULM--NAQ-----RLOVSEIEGIDRQGLKDOGLFIED-----ELI 195
DB 527 KYSELVQNHADILRLNAEYTKQVSMARQAQVD-LEREKKELEDLSERISDQQRTOQOL 585
QY 196 DLRLKYDGLVPRPRLIALEARAGLSLGS-----IGRLTADRASKAVGASDTQLK 245
DB 566 EYLESIKQELATNSQR--ELQVLYGSLSTSAQSEANMAAFALKEKRSIYSGAHHREEE 643
QY 246 VROIKOEFEQVYSOSITETRVRLAEVTEKEVVASDAQRIKIVSPVNGTAONLRFETGCA 305
DB 644 LSLRKE-----LQDTQLKLASTESMCQLAKDQKRMLLV---GSRKAEGVYIDA 691
QY 306 VVRAAE-PLVYDA-----PRD-----EAFVIOAHF 329
DB 692 INQLEEPPLISGASADHLLSTVTSISSCTEQLERKSWQYACLPDISGLLHSTITLHL 751
QY 330 OPTDVNVMGMVTEVRLPAFHSAGNPDPERHDPVAVADRISDPQKQ 376
DB 752 TS---DALHAGATTLRLAP-----PEP-----ADSLTEACKQ 780

```

RESULT 9

US-08-924-629C-6
 ; Sequence 6, Application US/08924629C
 ; Patent No. 6403082
 ; GENERAL INFORMATION:
 ; APPLICANT: Stiles, Michael E.
 ; APPLICANT: Vederas, John C.
 ; APPLICANT: van Belkum, Marius J.
 ; APPLICANT: Morobo, Randy W.
 ; APPLICANT: Morobo, Rodney J.
 ; APPLICANT: Greer, G. Gordon
 ; APPLICANT: McMullen, Lynn M.
 ; APPLICANT: Leisner, Jorgen J.
 ; APPLICANT: Poon, Alston
 ; APPLICANT: Franz, Charles M.A.P.
 ; TITLE OF INVENTION: No. 6403082 Antibacteriocins, Transport and Vector System and Method
 ; FILE REFERENCE: 660.000505
 ; CURRENT APPLICATION NUMBER: US/08/924,629C
 ; CURRENT FILING DATE: 1997-09-05
 ; PRIOR APPLICATION NUMBER: US 60/026,257
 ; PRIOR FILING DATE: 1996-09-05
 ; NUMBER OF SEQ ID NOS: 80
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 6
 ; LENGTH: 457
 ; TYPE: PRT
 ; ORGANISM: lcad;
 ; US-08-924-629C-6

Query Match 5.2%; Score 113.5; DB 4; Length 457;
 Best Local Similarity 18.4%; Pred. No. 0.017;
 Matches 85; Conservative 76; Mismatches 195; Indels 105; Gaps 16;

QY 7 QRPDNEQAVARIGYIALTFVGLGMAFAPLDSAVIANG-VYSAEVSODVOHLEGGM 65
 DB 14 QRRRNEFTLIIV--IFLLVFIILFLFAKREIVYKAGELIPAVLSDIGSTNNA 70
 QY 66 LAKLIVREGVYKAGVLEFDPTQANAAAGITRNOYVALKAMEARLAEEDQRPST--- 122
 DB 71 IDSNQLENNKYKKGDLVFTSGNEKISQLTQO--LNNLNDRLKSLDITYKOSIYNG 127
 QY 123 ---SEPAD-----LTSQRADPMVABAIADBOAQFTRRRTQIGOVNDLM 162
 DB 128 RSEGGTQFGYDLSFNQYMAQVDTLISEFNOOSSDKTADQQAHH-----QIDVL 178
 QY 163 ---NAORLYOSEIEGIDROTQGLKDLGFIED---ELIDRLKLYDKGLVPRRL 211
 DB 179 KOGOSKNNQOALANYQAILTSINSNTKPTNNPYQAIYDNYSAOLKSAQOTDDKQYKOTAL 238
 QY 212 L-----ALBARAGSLSGISGLTLADRSKAVQASDTQLAVRQIKQEFFEVSOSITET 264
 DB 239 SNVOQOQIDQLQTTSSSYDSQIAGIT--KSGPLSSSTLDRIDAKLQOOLASAKKEINDQ 295
 QY 265 RVRLAEVTEREVASDAQRIKIVSPVNGTAQ-----NLRFTEGAVVRAAEPLVDIAP 318
 DB 296 QOSIDELKAKKSSANEDQDVIVAKPEDGILHLATDKTKITFYFGKGTIAOYPRL---T 352
 QY 319 EDEAFVIOAHFOPTDV-----DNVHGMVTEVRLPAFHSAGNPDEPRHDV 364
 DB 353 QKTALNVEYVPASNIILKQROAIRFVANOVTKPLFLNGLTISISSA-----PI 403
 QY 365 AVADRISPOKQARLF-----LGIVRVYKOLPRLRGRTV 400
 DB 404 A-----SKESGFYKLAVATIOASKIDREQIKYGLNGRIT 436

RESULT 10
 ; US-09-134-001C-4920
 ; Sequence 4920, Application US/09134001C
 ; Patent No. 6380370
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 ; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: GTC-007
 ; CURRENT APPLICATION NUMBER: US/09/134,001C
 ; CURRENT FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/064,964
 ; PRIOR FILING DATE: 1997-11-08
 ; PRIOR APPLICATION NUMBER: US 60/055,779
 ; PRIOR FILING DATE: 1997-08-14
 ; NUMBER OF SEQ ID NOS: 5674
 ; SEQ ID NO 4920
 ; LENGTH: 531
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus epidermidis
 ; US-09-134-001C-4920

Query Match 5.2%; Score 112; DB 4; Length 531;
 Best Local Similarity 18.7%; Pred. No. 0.029;
 Matches 69; Conservative 67; Mismatches 153; Indels 80; Gaps 11;

QY 44 VIANGVYSAEVSODVOH--LEGMLAKILY---REGKVKAGVLEFDPTQ--ANAA 94
 DB 25 IILGVVGYIVARNILHOKQVQARQADDIYSANKKADNIKKEKLEAKEENQILKEQA 84
 QY 95 AGITRNOYVALKAMEARLAEEDQRPSTSPADITSQRADPMVABAIADBOAQFTERRQT 154
 DB 85 ENELRERREGELQREQETRLQKE--NLDRKSDDLDRDE---ILEQESKLEERQOQ 136
 QY 155 IOGOVDLMAORLYOSEIEGIDROTQ--GLKDOLGFIEDLIDRLKLYDKGLVPRRL 212
 DB 137 VDAKESSVQTLIMKHELEERISGLTOBEAVKBELOVEEEL----- 178
 QY 213 ALEARAGSLSGISGLTLADRSKAVQASDTQLKVRQIKQEFFEVSOSITETRVRLAEVT 272
 DB 179 -----SODIALILVKEKEKEAKEKRYDTAKEL---LATTV 209
 QY 273 EKEVYASDAQRIKIVSPVNGTAQNLRFTEGAVVRAEPL-----VDIAPDEARVIO 326
 DB 210 QRLAEHTTESTVSIVNLPDEMKGRIIGREGNRIRLETTLTGDLITDTPPE--AVIL 266
 QY 327 AHFOPTDVNHNQMVTEVRLPAFHSAGNPDEPRHDVAVADRISDPKQARLFLGIVRV 386
 DB 267 SCFPIREIRLRTALVNLVSDGRHHPRIEDMVAKARKEVDIIRDAQEATF-----EI 321
 QY 387 DYKOLPRL 395
 DB 322 NVHNMHPDL 330

RESULT 11
 ; US-09-085-199B-4
 ; Sequence 4, Application US/09085199B
 ; Patent No. 6235879
 ; GENERAL INFORMATION:
 ; APPLICANT: Hayden, Michael R.
 ; APPLICANT: Hackam, Abigail
 ; APPLICANT: Hug, A.H.M. Mahbubul
 ; APPLICANT: Chopra, Vikramjit Singh
 ; APPLICANT: Kalichman, Michael
 ; TITLE OF INVENTION: Apoptosis Modulators That Interact with the
 ; TITLE OF INVENTION: Huntington's Disease Gene
 ; NUMBER OF SEQUENCES: 44
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Oppe Dahl & Larson
 ; STREET: PO Box 5270
 ; CITY: Frisco
 ; STATE: CO
 ; COUNTRY: USA
 ; ZIP: 80443-5270
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Kb storage
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: MS DOS 5.0
 ; SOFTWARE: WordPerfect
 ; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/085,199B
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Larson, Marina T.
REGISTRATION NUMBER: 32038
REFERENCE/DOCKET NUMBER: UBC.P-0130US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (970) 668-2050
TELEFAX: (970) 668-2052
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 914
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: no
ORIGINAL SOURCE:
ORGANISM: human
FEATURE: Huntington-Interacting protein
US-09-085-199B-4

Query Match 5.2%; Score 112; DB 4; Length 914;
Best Local Similarity 20.1%; Pred. No. 0.067;
Matches 82; Conservative 72; Mismatches 131; Indels 122; Gaps 20;

QY 47 NGVSAEYSDVOHLE---GGMIAKILREGEKAVAGVLFELDPTQANAAAGITRNOYV 103
DB 243 NGVNDEKDHLEIRLYRELSGIAKQL---ENKKTESQRYVILQKHVSELEADLAEQCHL 299
QY 104 ALKAME--ARLIAERDQRPSTISFPADLTISQADPVAARAIDE---QAQETERR----- 152
DB 300 RQQAADDEFLRAELD-----ELRRQREDTEKAQRSLSEIRKQANQRYSKLKE 350
QY 153 ---QTIQGVNDL--NAQ-----RLQYQSEITGIDRQGLKQGLFIED-----ELI 195
DB 331 KYSELVQNHADLIRKNAETQYQVSMARQAYD-LEKEKKELEDSLESLSDQQRKTQEDL 409
QY 196 DLRLKYDGLVPRRLALLEARAGSLGS-----IGRLTADRSKAVAGASDTOLK 245
DB 410 EYLESKKELGTSOR--ELQVLOGLSLSTASQSEANMAAFELKEKERSLVSGAHHREBE 467
QY 246 VRQIKQEFFEYQSSITETRVRLAEVTEKEVVAASDAQRIKIKIVPVGTAQNLRFETEGA 305
DB 468 LSAIRKE-----LODTOLKLAESTEESMQLAKDKRMILLY-----GSRKAAEQVIDA 515
QY 306 VVRAAE-PLVNDIA-----PRD-----EAFVIOAHF 329
DB 516 LNLLEPPLISCAGADHLSTVTSISSCIQLEKSWSOYLACPDIDISGLLHSTITLHL 575
QY 330 QPTVDYNHMGVTEVRLPAFHSAGNPDPERHDPVAVADRIDSPQKQ 376
DB 576 TS---DATAHGATTCLRAP-----PEP-----ADSLTEACKQ 604

RESULT 12
US-09-599-652-2
Sequence 2, Application US/09599652
Patent No. RE37741

GENERAL INFORMATION:
APPLICANT: HOSTETTER, MARGARET K.
APPLICANT: GALE, CHERYL A.
APPLICANT: BENDEL, CATHERINE M.
APPLICANT: TAO, NIAN-JUN
APPLICANT: KENDRICK, KATHLEEN
TITLE OF INVENTION: CANDIDA ALBICANS GENE, INTEGRIN-LIKE
TITLE OF INVENTION: PROTEIN, ANTIBODIES, AND METHODS OF USE
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESS: MUEITING, RAASCH, GRABHARDT & SCHNAPPACH, P.A.
STREET: 119 NORTH FOURTH STREET, SUITE 203
CITY: MINNEAPOLIS
STATE: MINNESOTA

COUNTRY: USA
ZIP: 55401
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/599,652
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/642,846
FILING DATE: 03-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: MUEITING, ANN M.
REGISTRATION NUMBER: 33,977
REFERENCE/DOCKET NUMBER: 110,00280101
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-305-1217
TELEFAX: 612-305-1228
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1664 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-599-652-2

Query Match 5.0%; Score 109; DB 1; Length 1664;
Best Local Similarity 19.5%; Pred. No. 0.32;
Matches 80; Conservative 61; Mismatches 124; Indels 146; Gaps 20;

QY 92 NAAAGITRNOYALAKAMEKRLIAERDQRPSTISFPADLTISQADPVAARAIDEQAQETERR 151
DB 974 NAKKVTQDEYI---NAKLVDQKPKKNSI-----VTPDEPDEYEL 1010
QY 152 RQY-----IQGVNDLMAORLYQSEITGIDRQGLKQGLFIEDLIDLRK- 199
DB 1011 QQTASIHNAITDSSLYGRDSTIDMLRYLSD-ELKKRPYLLASDLRFMEBEVRLPSN 1069
QY 200 -----LYDKGLVPRRLALLEARAGSLGSIGRLTADRSKAVAGASDTOLKVRQI 249
DB 1070 SYLVHPGAGAATNSMLPEDEPDELINSPARVNSNDVY-----AIGNAST-ISPQL 1122
QY 250 KQEFFEQ--VQSITE-----TRVRLAEVTEKEVVAASDAQ 282
DB 1123 DMNFDQATIGQIKQIEOPASKSANTVRGDDDLASAPETPRTPRKESISSKPAKLSAS 1182
QY 283 KR---IKIVSPVNGTAQNLRFETEGAVVRAAEPLVNDIAPEDEAFVIOAHFQPTVDYDN--V 337
DB 1183 PRKSPKIGSPV-----RVTKKNGSIAGTPIPKAHKP-----KSPQGNELSNKV 1230
QY 338 HMGVTE-----VRLPAFHSAGN---PDP-----BRHD-- 362
DB 1231 RDGGISPSGSEHQHNPMSVSPSQYTDATSTVDENKMDVQHKPREKOKNHNHNHNH 1290
QY 363 -----PVAVADRIDSP--QKQARLF---LGIYRVVDYKQRLYLRGRTV 400
DB 1291 HHKQKTDIPGVVDDEIPVGLQERGLFEFRVLGINKINMLPDIINTH-KGRFT 1340

RESULT 13
US-08-642-846-2
Sequence 2, Application US/08642846
Patent No. 586151

GENERAL INFORMATION:
APPLICANT: HOSTETTER, MARGARET K.
APPLICANT: GALE, CHERYL A.
APPLICANT: BENDEL, CATHERINE M.
APPLICANT: TAO, NIAN-JUN
APPLICANT: KENDRICK, KATHLEEN

TITLE OF INVENTION: CANDIDA ALBICANS GENE, INTEGRIN-LIKE
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESSES:
ADDRESS: MUEITING, RAASCH, GEBHARDT & SCHWAPPACH, P. A.
STREET: 119 NORTH FOURTH STREET, SUITE 203
CITY: MINNEAPOLIS
STATE: MINNESOTA
COUNTRY: USA
ZIP: 55401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/642,846
FILING DATE: 03-MAY-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: MUEITING, ANN M.
REGISTRATION NUMBER: 33,977
REFERENCE/DOCKET NUMBER: 110.00280101
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-305-1217
TELEFAX: 612-305-1228
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1664 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-642-846-2

Query Match 5.0%; Score 109; DB 2; Length 1664;
Best Local Similarity 19.5%; Pred. No. 0.32;
Matches 80; Conservative 61; Mismatches 124; Indels 146; Gaps 20;

QY 92 NAAAGITRNOVALKAMARLARLAERDORPSISFPADLTSGRADPVVARAIADQOFTER 151
DB 974 NAKKGVTODEYI-----NAKLVDOKPKKNSI-----VTDPEDEYIEL 1010

QY 152 ROT-----IOGVDMNARLOYOSEIEGIDRQGLKQDLGFIDELELDLKK- 199
DB 1011 QOTASINATIDSSITGPRDSISTDMLPYLSD-ELKKPTALLSADRLFMEOEVNRLKSN 1069

QY 200 -----LYDKGLVPRRLALLAERAGSLSGISGRLTADRKAAGASDQTLKVRQI 249
DB 1070 SVLVHPGAGATNSMRLPEPFELINSPARVNSNSDNV-----AISGNAST-ISFNOL 1122

QY 250 KOEFFEQ--VSOSITE-----TVRLAEVTEKEVVASDAQ 282
DB 1123 DMNDDQATIOKIOEORPASKANTVRGDDGLASARETPRTPTKESISSKRAKLSSAS 1182

QY 283 KR---IKIVSVNGTAQNLREFTEGAVVRAERLVDIAREDEAFVIANFOPTVDN--V 337
DB 1183 PRKSPIKIGSV-----RVTKKNGSIAGIEPIPKATHKP-----KKSFGNGEISNKKV 1230

QY 338 HMGWTE-----VRLPFSAGN---PDP-----ERND-- 362
DB 1231 RDGGISPSGSEHQHNPSVSPSQYTDATSTVDENKDVONKPREKOKKNNHNNHNN 1290

QY 363 -----PVAVADRISDP--OKQARLF---LGIVRVDKQLRPRLKGRVT 400
DB 1291 HHKOKTDIRGVVDEIRPVGLQERGKILFFRVLGIKINIMLPDINTN-KGRFT 1340

RESULT 14
US-09-264-604-2
Sequence 2, Application US/09264604
Patent No. 6346411
GENERAL INFORMATION:

APPLICANT: HOSTETTER, MARGARET K.
APPLICANT: GALE, CHERYL A.
APPLICANT: BENDEL, CATHERINE M.
APPLICANT: TAO, NIAN-JUN
APPLICANT: KENDRICK, KATHLEEN
TITLE OF INVENTION: CANDIDA ALBICANS GENE, INTEGRIN-LIKE
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESSES:
ADDRESS: MUEITING, RAASCH, GEBHARDT & SCHWAPPACH, P. A.
STREET: 119 NORTH FOURTH STREET, SUITE 203
CITY: MINNEAPOLIS
STATE: MINNESOTA
COUNTRY: USA
ZIP: 55401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/264,604
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/642,846
FILING DATE: 03-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: MUEITING, ANN M.
REGISTRATION NUMBER: 33,977
REFERENCE/DOCKET NUMBER: 110.00280101
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-305-1217
TELEFAX: 612-305-1228
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1664 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-264-604-2

Query Match 5.0%; Score 109; DB 4; Length 1664;
Best Local Similarity 19.5%; Pred. No. 0.32;
Matches 80; Conservative 61; Mismatches 124; Indels 146; Gaps 20;

QY 92 NAAAGITRNOVALKAMARLARLAERDORPSISFPADLTSGRADPVVARAIADQOFTER 151
DB 974 NAKKGVTODEYI-----NAKLVDOKPKKNSI-----VTDPEDEYIEL 1010

QY 152 ROT-----IOGVDMNARLOYOSEIEGIDRQGLKQDLGFIDELELDLKK- 199
DB 1011 QOTASINATIDSSITGPRDSISTDMLPYLSD-ELKKPTALLSADRLFMEOEVNRLKSN 1069

QY 200 -----LYDKGLVPRRLALLAERAGSLSGISGRLTADRKAAGASDQTLKVRQI 249
DB 1070 SVLVHPGAGATNSMRLPEPFELINSPARVNSNSDNV-----AISGNAST-ISFNOL 1122

QY 250 KOEFFEQ--VSOSITE-----TVRLAEVTEKEVVASDAQ 282
DB 1123 DMNDDQATIOKIOEORPASKANTVRGDDGLASARETPRTPTKESISSKRAKLSSAS 1182

QY 283 KR---IKIVSVNGTAQNLREFTEGAVVRAERLVDIAREDEAFVIANFOPTVDN--V 337
DB 1183 PRKSPIKIGSV-----RVTKKNGSIAGIEPIPKATHKP-----KKSFGNGEISNKKV 1230

QY 338 HMGWTE-----VRLPFSAGN---PDP-----ERND-- 362
DB 1231 RDGGISPSGSEHQHNPSVSPSQYTDATSTVDENKDVONKPREKOKKNNHNNHNN 1290

QY 363 -----PVAVADRISDP--OKQARLF---LGIVRVDKQLRPRLKGRVT 400

Db 1291 HHKORTDIPGVDDDEIPVGLQERKLFRRVLGINKINLPDINTH-KGRFT 1340

RESULT 15

US-08-816-693A-51
Sequence 51, Application US/08816693A
Patent No. 5874241

GENERAL INFORMATION:

APPLICANT: Takahashi, Joseph S

APPLICANT: Turek, Fred W

APPLICANT: Pinto, Lawrence H

TITLE OF INVENTION: Clock Gene and Gene Product

NUMBER OF SEQUENCES: 53

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dressler, Rockey, Milnamow & Katz

STREET: Two Prudential Plaza, Suite 4700

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/816,693A

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: No. 5874241thrup, Thomas E

REGISTRATION NUMBER: 33,268

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-616-5460

TELEFAX: 312-616-5460

INFORMATION FOR SEQ ID NO: 51:

SEQUENCE CHARACTERISTICS:

LENGTH: 747 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-816-693A-51

Query Match 4.9%; Score 107; DB 2; Length 747;

Best Local Similarity 19.6%; Pred. No. 0.15; Matches 72; Conservative 55; Mismatches 101; Indels 140; Gaps 17;

QY 86 LDPTQANA---AGIT-----RNOYALK---AMEARLLAERDQPSISFP--- 125
DB 378 LEPRQFNLGDGASGLSPSSASRSKSHSTHAMESEPISTPKLMAESTALPRATLPOL 437
QY 126 -----ADLTQORADPVAARAIDE-QAOFTEERRQTIQGVDDIMNAQRIQ- 168
DB 438 PVGLSOAATMPLSSCDITQQLQPLQSPAPQPSAOPS-MFQTKDQLE---QRTRI 492
QY 169 YQSEIEGIDRQOGKLDGFIEDDELIDRKLYDKGIVRPRRLALEARASLSGSIGRL 228
DB 493 LDANIRWQOEELHKIQEQLCLVQDSNVQ-----MFLQDPVSLSES----- 533
QY 229 TADRSKAVOGASDTQLKVKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 282
DB 534 -----STORPAQQLQQAQAAPOLVQLOGQISTQVTOHLARESSVYISGPKPKMSSQLS 587
QY 283 KRIKIVSPVNGTAQNRFTBGAIVRAAEPLVDI---APEDEAFVIAQAHFOPTDVQNVHM 339
DB 588 GRSSSLSPFSST-----LPPLTTPASTPQD----- 613
QY 340 GAVTEVRLPAFHSAGNPDPERHDPVAVADRIDSPQKQARFLGIVRVQKOLPHLGRV 399
DB 614 -----SQCOPSPDFHD-----RQLRLLS-----QPIQPMKPGSC 643
QY 400 TAGMPAQV 407

Db 644 DARQPSV 651

Search completed: April 22, 2003, 12:33:14
Job time : 23 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:27:47 ; Search time 35 Seconds

(without alignments)
2560.871 Million cell updates/sec

Title: US-09-913-414-5

Perfect score: 2165
Sequence: 1 MKPKRQRPDNDQAVARIG.....VLQYLFSPRLDRLTTRREE 435

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protist:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaeal:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1984.5	91.7	436	16	085351 caulobacter
2	1972.5	91.1	436	2	09RMW9 caulobacter
3	606	28.0	436	16	092NX2 rhizobium m
4	589	27.2	436	2	09X6N7 rhizobium l
5	589	27.2	474	16	08UAI1 agrobacteri
6	583	26.9	439	16	033679 rhizobium m
7	568	26.2	436	16	08U7N5 agrobacteri
8	564	26.1	443	16	09HYJ9 pseudomonas
9	560.5	25.9	434	16	098U69 rhizobium l
10	543.5	25.1	437	16	P96438 rhizobium m
11	542.5	25.0	437	16	08UB82 rhizobium m
12	540.5	25.0	443	2	054417 serratia ma
13	536.5	24.8	473	2	005198 rhizobium m
14	530.5	24.5	511	16	092IT2 rickettsia
15	522	24.1	443	2	09XB63 erwinia amy
16	521.5	24.1	443	2	054457 serratia ma

17	519.5	24.0	511	16	09ZDL4 rickettsia
18	512	23.6	435	2	005694 rhizobium l
19	510	23.6	433	2	09XG93 pseudomonas
20	509	23.5	437	2	067994 pseudomonas
21	508.5	23.5	443	2	08VLR7 yersinia ru
22	502.5	23.2	447	16	0988E5 rhizobium l
23	499	23.0	435	2	09ECN8 pseudomonas
24	495	22.9	444	2	09ZNT8 pseudomonas
25	493	22.8	437	2	087809 pseudomonas
26	489.5	22.6	427	2	09ZIG7 campylobact
27	482.5	22.3	427	2	09R8D9 campylobact
28	479.5	22.1	443	2	09RHT1 pseudomonas
29	479	22.1	439	2	09R9H4 pseudomonas
30	477	22.0	438	2	09KGS5 pseudomonas
31	476.5	22.0	442	16	08ZAG2 yersinia pe
32	476.5	22.0	450	16	09ZLS8 rhizobium m
33	471.5	21.8	437	2	057387 serratia ma
34	467.5	21.6	440	16	098140 rhizobium l
35	467.5	21.6	441	2	085376 proteus mir
36	438.5	20.3	497	16	08UAA0 agrobacteri
37	436	20.1	433	2	052859 rhizobium l
38	393	18.2	358	16	08XQ99 rickettsia s
39	375	17.3	387	16	08ZMY7 salmonella
40	368.5	17.0	387	16	08Z4H8 salmonella
41	368	17.0	481	16	09KRM3 vibrio chol
42	367	17.0	395	16	0912M0 pseudomonas
43	329	15.2	475	16	09JY53 neisseria m
44	326.5	15.1	475	2	09X580 neisseria m
45	326.5	15.1	475	2	09JPK4 neisseria m

ALIGNMENTS

RESULT 1
ID 085351 PRELIMINARY; PRT; 436 AA.
AC 085351;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Membrane forming UNIT (R5AA secretion system, membrane protein
DE R5AE).
GN R5AE OR CCI1009.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCB1_TaxID=155892;
OX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15, AND NA1000;
RX MEDLINE=98292737; PubMed=9620954;
RA Avram P., Smit J.;
RT "The Caulobacter crescentus paracrystalline S-layer protein is
RT secreted by an ABC transporter (type I) secretion apparatus.";
RT J. Bacteriol. 180:3062-3069(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15, AND NA1000;
RA Avram P., Smit J.K.;
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15, AND NA1000;
RA Avram P.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
[4]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nieman W.C., Feldblum T.V., Laud M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,

RA Deboy R.T., Dodson R.J., Durkin A.S., Gwin M.L., Haft D.H.,
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
 RA Kutterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.,
 RT "Complete genome sequence of *Caulobacter crescentus*,"
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 DR EMBL: AF062345; AAC38667.2;
 DR EMBL: AF062345; AAC38667.2;
 DR TIGR: CC1009;
 DR InterPro: IPR002215; HLYD.
 DR InterPro: IPR003997; REXD.
 DR Pfam: PF00529; HLYD.1
 DR PRINTS: PR01490; RYTXOXND.
 DR PROSITE: PS00543; HLYD_FAMILY; UNKNOWN.1.
 KW Complete proteome.
 SQ SEQUENCE 436 AA; 48389 MM; 90E32A834673C9DC CRC64;

Query Match 91.7%; Score 1984.5; DB 16; Length 436;
 Best Local Similarity 93.6%; Pred. No. 1,3e-116;
 Matches 408; Conservative 3; Mismatches 24; Indels 1; Gaps 1;

QY 1 MKPRIQPTDNFQAVARIIGYIIALTFVGLGMAAFAPLDSAVIANGVSAEVS-ODVQ 59
 DB 1 MKPRIQPTDNFQAVARIIGYIIALTFVGLGMAAFAPLDSAVIANGVSAEGRKTYQ 60
 QY 60 HLEGMLAKILVREGEKVKAGOVLELDPQTQANAAAGITRNQYALKAMEARLLAERDOR 119
 DB 61 HLEGMLAKILVREGEKVKAGOVLELDPQTQANAAAGITRNQYALKAMEARLLAERDOR 120
 QY 120 PSISFPADLTSGRADPMVARIADQAOFTERRQTIQGVDMNORLOYOSEIEGIDRQ 179
 DB 121 PSISFPADLTSGRADPMVARIADQAOFTERRQTIQGVDMNORLOYOSEIEGIDRQ 180
 QY 180 TQGLKDQGFIEDELIDRLKLYDKGLVPRRLALEARQASLSGIGRLTADRASKAVOGA 239
 DB 181 TQGLKDQGFIEDELIDRLKLYDKGLVPRRLALEARQASLSGIGRLTADRASKAVOGA 240
 QY 240 SDTQKVKQIKOEFEQVQSSTTERVRVLAETKEEVASDAQKRIKIVSPVNGTAQMLR 299
 DB 241 SDTQKVKQIKOEFEQVQSSTTERVRVLAETKEEVASDAQKRIKIVSPVNGTAQMLR 300
 QY 300 FETEGAVVRAAEPLVDIAPEDAEFYIOAHFOPTVDVNMGMVTEVRLLPAFHSAGNPDE 359
 DB 301 FETEGAVVRAAEPLVDIAPEDAEFYIOAHFOPTVDVNMGMVTEVRLLPAFHSAGNPDE 360
 QY 360 RHDPVAVADRISDPQKARLELGIYRVYKQLPRLRGRTVAGMPAOYIVPTGERTVLYQ 419
 DB 361 GTIOSLSODRISDPQKARLELGIYRVYKQLPRLRGRTVAGMPAOYIVPTGERTVLYQ 420
 QY 420 LFSPLRDLRTTMRRE 435
 DB 421 LFSPLRDLRTTMRRE 436

RESULT 2

Q9RMW9 PRELIMINARY; PRT; 436 AA.
 AC Q9RMW9;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE Membrane forming unit.
 GN R5AE.
 OS *Caulobacter crescentus*.
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
 OC Caulobacter.
 OX NCBI_TaxID=155892;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=J54050;
 RT Bingle W.H., Avram P.A., Nornellini J.F., Smit J.K.;
 RT "The secretion signal of the *Caulobacter crescentus* S-layer protein is
 located within the C-terminal 82 amino acids of the molecule."

RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF193064; AAF07963.1;
 DR InterPro: IPR002215; HLYD.
 DR InterPro: IPR003997; REXD.
 DR Pfam: PF00529; HLYD.1
 DR PRINTS: PR01490; RYTXOXND.
 DR PROSITE: PS00543; HLYD_FAMILY; UNKNOWN.1.
 SQ SEQUENCE 436 AA; 48443 MM; F9C4C680F3D96563 CRC64;

Query Match 91.1%; Score 1972.5; DB 2; Length 436;
 Best Local Similarity 93.1%; Pred. No. 7.4e-116;
 Matches 406; Conservative 3; Mismatches 26; Indels 1; Gaps 1;

QY 1 MKPRIQPTDNFQAVARIIGYIIALTFVGLGMAAFAPLDSAVIANGVSAEVS-ODVQ 59
 DB 1 MKPRIQPTDNFQAVARIIGYIIALTFVGLGMAAFAPLDSAVIANGVSAEGRKTYQ 60
 QY 60 HLEGMLAKILVREGEKVKAGOVLELDPQTQANAAAGITRNQYALKAMEARLLAERDOR 119
 DB 61 HLEGMLAKILVREGEKVKAGOVLELDPQTQANAAAGITRNQYALKAMEARLLAERDOR 120
 QY 120 PSISFPADLTSGRADPMVARIADQAOFTERRQTIQGVDMNORLOYOSEIEGIDRQ 179
 DB 121 PSISFPADLTSGRADPMVARIADQAOFTERRQTIQGVDMNORLOYOSEIEGIDRQ 180
 QY 180 TQGLKDQGFIEDELIDRLKLYDKGLVPRRLALEARQASLSGIGRLTADRASKAVOGA 239
 DB 181 TQGLKDQGFIEDELIDRLKLYDKGLVPRRLALEARQASLSGIGRLTADRASKAVOGA 240
 QY 240 SDTQKVKQIKOEFEQVQSSTTERVRVLAETKEEVASDAQKRIKIVSPVNGTAQMLR 299
 DB 241 SDTQKVKQIKOEFEQVQSSTTERVRVLAETKEEVASDAQKRIKIVSPVNGTAQMLR 300
 QY 300 FETEGAVVRAAEPLVDIAPEDAEFYIOAHFOPTVDVNMGMVTEVRLLPAFHSAGNPDE 359
 DB 301 FETEGAVVRAAEPLVDIAPEDAEFYIOAHFOPTVDVNMGMVTEVRLLPAFHSAGNPDE 360
 QY 360 RHDPVAVADRISDPQKARLELGIYRVYKQLPRLRGRTVAGMPAOYIVPTGERTVLYQ 419
 DB 361 GTIOSLSODRISDPQKARLELGIYRVYKQLPRLRGRTVAGMPAOYIVPTGERTVLYQ 420
 QY 420 LFSPLRDLRTTMRRE 435
 DB 421 LFSPLRDLRTTMRRE 436

RESULT 3

Q92NX2 PRELIMINARY; PRT; 436 AA.
 AC Q92NX2;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Putative toxin secretion transmembrane protein.
 GN R02036 OR SMC04208.
 OS *Rhizobium melioli* (*Sinorhizobium melioli*).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; *Sinorhizobium*.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RC MEDLINE=21396507; PubMed-11481430;
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
 RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masny D.,
 RA Pohl T., Portetelle D., Puehler A., Punnelle B., Rameberger U.,
 RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.;
 RT "Analysis of the chromosome sequence of the legume symbiont
Sinorhizobium melioli strain 1021."
 RT Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
 DR EMBL: AL591789; CAC46615.1;
 DR InterPro: IPR002215; HLYD.

ID	09X6N7	PRELIMINARY:	PRT:	436	AA:
AC	09X6N7;				
DT	01-NOV-1999 (TREMBLrel. 12, Created)				
DT	01-NOV-1999 (TREMBLrel. 12, Last sequence update)				
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)				
DE	Rhizobium secretion protein RspE.				
GN	RspE.				
OS	Rhizobium leguminosarum (biovar trifolii).				
OC	Plasmid pRhe162Y10C.				
OC	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;				
OC	Rhizobiaceae; Rhizobium.				
NCBI_TaxID=386;					
11]					
SEQUENCE FROM N.A.					
STRAIN-162Y10;					
MEDLINE-99318631; PubMed-10388672;					
RA	Oresnik I.J., Twelker S., Hynes M.F.;				
RT	"Cloning and characterization of a Rhizobium leguminosarum gene				
RT	encoding a bacteriocin with similarities to RTX toxins.";				
Appl. Environ. Microbiol. 65:2833-2840(1999).					
12]					
SEQUENCE FROM N.A.					
STRAIN-162Y10;					
MEDLINE-21360352; PubMed-11467725;					
RA	Venter A.P., Twelker S., Oresnik I.J., Hynes M.F.;				
RT	"Analysis of the genetic region encoding a novel rhizobium from				
RT	Rhizobium leguminosarum by virulence strain 306.";				
Can. J. Microbiol. 47:495-502(2001).					
EMBL; AF141933; AAD4891.1; -					
InterPro; IPR002215; HL5D.					

RESULT 5			
OB00A11			
ID	OB00A11	PRELIMINARY;	PRT; 474 AA.
AC	OB00A11;		
DT	01-JUN-2002 (TEMBUREL, 21, Created)		
DT	01-JUN-2002 (TEMBUREL, 21, last sequence update)		
DT	01-JUN-2002 (TEMBUREL, 21, last annotation update)		
DE	Secretion protein, HLYD family.		
GN	PSSE OR ATU0382 OR AGR_L12881.		
OS	Agrobacterium tumefaciens (strain C58 / ATCC 33970).		
OC	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;		
OC	Rhizobiaceae; Rhizobium.		
OX	NCBI_TaxID=176299;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE:21608550; PubMed:11743193;		
RA	Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,		
RA	Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,		
RA	Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,		
RA	Chapman P., Clendenenling J., Deatherage G., Gilllet W., Grant C.,		
RA	Kutayavin T., Levy R., Li M.-Y., McClelland E., Palmeri A.,		
RA	Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,		
RA	Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan M., Perry M.,		
RA	Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,		
RA	Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,		
RT	Nester E.W.;		
RT	"The genome of the natural genetic engineer Agrobacterium tumefaciens		
RT	C58."		
RL	Science 294:2317-2323(2001).		
RN	[2]		

RP SEQUENCE FROM N.A.
RA York G.M., Walker G.C.;
RT "The *Rhizobium meliloti* exoX gene and prsD/prsE/exsH genes encode
RT components of independent degradative pathways which contribute to
RT production of low-molecular-weight succinoglyca.";

RX MEDLINE=21608550; PubMed=11743193;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eissen J.A., Karp P.D., Bovee D. Sr.,
 RA Chapman P., Clendenning J., Deatherage G., Gillitt W., Grant C.,

RA Kutyavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
 RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nester E.W.;
 RT "The genome of the natural genetic engineer *Agrobacterium tumefaciens*
 C58.";
 RL Science 294:2317-2323(2001).
 RP [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
 RA Houliet B., Goldman B.S., Cao Y., Askenazi M., Hailing C., Mullin L.,
 RA Hummel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
 RA Wollam C., Allinger M., Doughy D., Scott C., Iappas C., Markelz B.,
 RA Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,
 RA Clelo C., Slater S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent
 RT *Agrobacterium tumefaciens* C58.";
 RL Science 294:2323-2328(2001).
 DR EMBL; AE009370; AAI45208.1; ALT_INIT.
 DR EMBL; AE008244; AAK89025.1; -.
 KW Complete proteome.
 SQ SEQUENCE 436 AA; 48511 MW; 9C47850D301E9863 CRC64;

Query Match 26.2%; Score 568; DB 16; Length 436;
 Best Local Similarity 31.5%; Pred. No. 7.9e-28;
 Matches 138; Conservative 79; Mismatches 20; Indels 20; Gaps 5;

OY 11 DNFOAVARI-----GYGIALTFVGLGMAAFPLDSAVIANGVSAEVS-ODVOHL 61
 DB 4 DEMQALRSIRSHLVGLGFLTLGV-FGMAVGTETIYAGVLAQSLVETSLKKVQHP 62
 OY 62 EGMGLAKILVREGGEYKAGVLFELDPYQANAAGITRNOYVALKAMEARLLAEFDQRP 121
 DB 63 VGVVSELMVRDGDVKKAGDVYMRIDATMTTRANLAIVKSLDQFTARRKRLGEGEDRAAS 122
 OY 122 ISFPADLTQSORDPVARIAIDEOAQOETERRQTIGOVDMNAORLOYSSEIGIDRQT 181
 DB 123 VVFPQSLDRAGDAEVLAMNAEOQLYEDRAVRRESKRLQORVROLDELISGLEAERA 182
 OY 182 GIKDQGLFEDLEIDLRLKYDKGLVPRPLALAEARAGSLSSIGRLTADRKAQVQASD 241
 DB 183 ANVRQGWDEELIFRSIQERGLDLSKSLTLEQATDIDIDIGRLAGINGIAKISE 242
 OY 242 TQLKROIKQEFEEYQSITETRYRLAEVTEKEVYASDAQRKITVSPVNGTAONLRF 301
 DB 243 TALQILQIDQWSEVGSDLREMDARIGEYERRVAAEQDLKRVILAPQDGVHQLAVH 302
 OY 302 TEGAVVRAAEPLVDIAPEDEAFVIOAHFOPTVDVNVHMGVTEVRLPAFHSAGNDPDERH 361
 DB 303 TVGGVIAAGEQIMLVPEVDKLVFAKVAPODIDQITFGQVNLNLSAFNQSTTP----- 357
 OY 352 DEVAIVADRI-----DPQKARLFLGIVRVYDKQLPRHLRGVTAQMPRAQVITPTGERTV 416
 DB 358 EITGTERISADVVDQFGASGYLVAVATSEQDLKRLCEFSLMGMPVEARITTGERSV 417
 OY 417 LQYLFSPRLDITRTMRE 434
 DB 418 LSYFLKPLLDQANRTFRQ 435

RESULT 8
 O9HYJ9 PRELIMINARY; PRT; 443 AA.
 AC O9HYJ9;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Metalloprotease secretion protein.
 GN HSB OR PA3405.
 OS *Pseudomonas aeruginosa*.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Steyer C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Laddig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of *Pseudomonas aeruginosa* PA01, an
 RT opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 DR EMBL; AE004761; AAG06793.1; -.
 DR InterPro; IPR002215; HLYD.
 DR InterPro; IPR003997; RTXD.
 DR Pfam; PF00529; HLYD; 1.
 DR PRINTS; PR01490; RTXTOXIND.
 KW Protease; Complete proteome.
 SQ SEQUENCE 443 AA; 48776 MW; 06CCD3E9DB7092A7 CRC64;

Query Match 26.1%; Score 564; DB 16; Length 443;
 Best Local Similarity 32.5%; Pred. No. 1.4e-27;
 Matches 138; Conservative 87; Mismatches 189; Indels 10; Gaps 4;

OY 17 ANIGYIALTFVGLGMAAFPLDSAVIANGVSAEVSQD-VQHLGGLAKILVREG 75
 DB 25 SRLGMLVLLGFGVGLMAGLAPLDKGVSGTVMAGSRKAVOHPGTGLVHRHREGE 84
 OY 76 KYKAGVLFELDPYQANAAGITRNOYVALKAMEARLLAEFDQRPISFPALTSQRADP 135
 DB 85 RVKAGVLEMDATQARRADGLFRQYLAALSLARLSERREKARIEFPALLA-LDDP 143
 OY 136 MYARIADEQAQOETERRQTIGOVDMNAORLOYSSEIGIDRQTGKQGLFTEDELI 195
 DB 144 RLPTLEQQRQHLDRRRRLRLDELGLAETVAGSQALDGLALRSKQQRALAEQDLR 203
 OY 196 DLRLTYDKGLVPRPLALAEARA-----GSLSSIGRLTADRKAQVQASDQLKROIKQ 251
 DB 204 GRLQASEGYVRNRLDSERLLAQVNGEIALDGLSLSTRQIL-----ELRLMAQRAE 259
 OY 252 EFFEVSQSITETRYRLAEVTEKEVYASDAQRKITVSPVNGTAONLRFEGAVVRAE 311
 DB 260 KQEEVRSALDAQVRAELRLKRLASARPLANSEYRAVAGLVYQGVFTGGVYIARQ 319
 OY 312 PLVDIAPEDEAFVIOAHFOPTVDVNVHMGVTEVRLPAFHSAGNDPDERHDPVAVADRIS 371
 DB 320 QLMETLPERQPLVDARLPEVVDKRVGLPVELMFSAFNQSTTPRVEGEVTVSADRLL 379
 OY 372 DPQKARLFLGIVRVYDKQLPRHLRGVTAQMPRAQVITPTGERTV LQYLFSPRLDITRT 431
 DB 380 DERSEAPYRVRIKVGEEVRLAGLEIRGMPVEAFVRSERSILNLEKPLADRTILA 439
 OY 432 MREE 435
 DB 440 LGEE 443

RESULT 9
 O9BLG9 PRELIMINARY; PRT; 434 AA.
 AC O9BLG9;
 DT 01-OCT-2001 (TREMBLrel. 18, Created)
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Rhizobium secretion protein, RspE.
 GN ML1026.
 OS *Rhizobium loti* (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Phyllobacteriaceae; Mesorhizobium.

ID	Accession	Query Match	Score	DB	Length
AC	080B22	PRELIMINARY	PRT	437	AA
DT	01-JUN-2002	(TREMblrel, 21, Created)			
DT	01-JUN-2002	(TREMblrel, 21, Last sequence update)			
OX	NCHI_TaxID=382;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=RCR2011 / SUB47;				
RX	MEDLINE=97175570; PubMed=9023225;				
RA	Becker A., Rueberg S., Kuester H.,				
RA	Ivashina T., Cheng H., Walker G.C., Puhler A.;				
RT	"The 33-kilobase exp gene cluster of Rhizobium meliloti directing the				
RT	biosynthesis of galactoglucan: genetic organization and properties of				
RT	the encoded gene products."				
RL	J. Bacteriol. 179:1375-1384(1997).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=1021; PLASMID=PSYMB (MEGAPLASMID 2);				
RX	MEDLINE=21396508; PubMed=11481431;				
RA	Finan T.M., Weidner S., Wong K., Buhmester J., Chain P.,				
RA	Vornheller F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,				
RA	Golding B., Puhler A.;				
RT	"The complete sequence of the 1,683-kb psymb megaplasmid from the N2-				
RT	fixing endosymbiont Sinorhizobium meliloti."				
RL	Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).				
DR	EMBL: Z79692; CAB01939.1; -				
DR	EMBL: AL603645; CAC49291.1; -				
DR	InterPro: IPR002215; HLYD.				
DR	InterPro: IPR003997; REXD.				
DR	Pfam: PF00529; HLYD; 1.				
DR	PRINTS: PR01490; RXTTOXIND.				
KW	Hypothetical protein; Plasmid; Complete proteome.				
SO	SEQUENCE 473 AA; 51479 MW; C0C3B16BD9C23D85 CnC64;				
Query Match	25.1%; Score 543.5; DB 16; Length 473;				
Best Local Similarity	28.8%; Pred. No. 3e-26;				
Matches 126; Conservative 97; Mismatches 207; Indels 7; Gaps 3;					
OY	4 PKIQPTNFQVAVR-----IGYIATLFFVGLGMAAFAPLDSAVIANVSAEYS-QDV 58				
DB	39 PEPLDSEDNTHSPIRKLRYLAGITLLVAFGFGFNAFSTELSSASVTSOTIYVDSKRTYV 98				
OY	59 OHLESGMLAKLIVREGGEKVGQVLFELDPTQANAAAGITRNQYVALKAMEARLLAERDQ 118				
DB	99 SHFEGGVIGRLIYVQGDHVAFCQPLMKLEDTFARSDDLALQSRVGLIAKLRLAERAG 158				
OY	119 RSTSEFPDLTQSQRADPWARAIADEQAQFTERRQTICQGVINDMAQRICQYSELEGIDR 178				
DB	159 LHAVFSPDLVSE--GEAAADVATAEKAFFERSEKESKRLIQRKTEIYSEKASGLTA 216				
OY	179 QTOGKLDLGFEDDELIDLKLVDYDGLVPRPRYLALEAAGSLSSIGRTLADRASKAVOG 238				
DB	217 QLOATDRQIELMNEQRATATILEKAFQRSKLIEIDANLSLAAITRGLADKKAQAEKA 276				
OY	239 ASDTOLKVRQIKQEFEEVOYSOSITETRVRLAEVTEKEVVASDAQRIKIVSVNGTAOML 298				
DB	277 MAGALALALGIESDQSESLAGEITFARLELADEVORITAAEVLRLERAFRAQDAIVANI 336				
OY	299 RFEETGAVVRAAEPLVYDIAPREDEAVIQAQFOPTVDNVHMGMAEVRIPARFSAGNPD 358				
DB	337 QLRTEGSAVTPQQLDIDYREDEPLLVENHVSITRDIIDSTISSQIRKLTAINQSHPL 396				
OY	359 ERHDEVAADRIISDPQOKARLFGLIVRVQDKPLPHLRGRYTAGMPAOYIVPTGERTYIQ 418				
DB	397 EKGVTYIADQSDVDEKSNVAEVFAAEVYTESLIANPDIRLYPGMPAEVLIVHKSSAID 456				
OY	419 YLFSPLRDTLRTTME 435				
DB	457 YLVAFVSDFNRAFRED 473				

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE HlyD family secretion protein.
 GN PRSE OR ATU2706 OR AGR_C_4904.
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743193;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boyce D. Sr.,
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
 RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
 RA Raymond C., Rouse G., Senphimachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nester E.W.;
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
 C58.";
 RL Science 294:2317-2323(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
 RA Ouyollo B., Goldman B.S., Cao Y., Askenzi M., Halling C., Mullin L.,
 RA Houmel K., Gordon J., Vaudin M., Tatchouk O., Epp A., Liu F.,
 RA Wollam C., Allinger M., Dougherty D., Scott C., Lapps C., Markelz B.,
 RA Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
 RA Cielo C., Slater S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent
 Agrobacterium tumefaciens C58.";
 RL Science 294:2323-2328(2001).
 DR EMBL; AE009216; AAL43687.1; -
 DR EMBL; AE008183; AAK88425.1; -
 KW Complete proteome.
 SQ SEQUENCE 437 AA; 4773 MW; 67479F1D40122D04 CRC64;

Query Match 25.1%; Score 542.5; DB 16; Length 437;
 Best Local Similarity 30.8%; Pred. No. 3.2e-26;
 Matches 130; Conservative 83; Mismatches 204; Indels 5; Gaps 2;
 Oy 15 AAVAGYGIATLFGGLGMAFAPLDSAV-ANGVSAE-VSODVOHLEGMALIKLVREGK 73
 Db 20 AVVVLGLGLL-----VGMGMAFAPLDSAV-ANGVSAE-VSODVOHLEGMALIKLVREGK 75
 Oy 74 GERVAGVLFELDPTQANAAGITRNOYVALKAMEARLLAERDORPSISPPADLTSGORA 133
 Db 76 GGRVAGVLFELDPTQANAAGITRNOYVALKAMEARLLAERDORPSISPPADLTSGORA 135
 Oy 134 DPMVARLADDEQAOTERRQTOGVDLMNORLOQOSIEBIDQOTGLKQOLGFIETDE 193
 Db 136 SKSARTFIDSEONLNSRNALIGMKKOLATRRKQLADARGLDQVATENELIAVYKED 195
 Oy 194 LIDLRLKLDKGVPRPRLLALEARAGSLGSGRLTADRSKAVOGASPDQLVROIKOFE 253
 Db 196 VSKTEBLAKKGLVLTORLNLKRLQSLNLEGGOGYIARAQYVGLSELDQLDQDEBR 255
 Oy 254 FEQVSOSITETRVRLAEVTEKEEVASDAQRIKIVSPVNGTAQNLRFETGAVVRAEPL 313
 Db 256 KSEVTKDLTSIBATVAEEERLAATRDQDLDRISPIGRITQYVSHNINCVIDPGEVL 315
 Oy 314 VDIAPDEDFVYQAHFOPIDVDVNHMGVTEVRLPAFHSGAPDPERHDPVAVADRISP 373
 Db 316 MLVVPDKDLDLEANNITPDDIDQIVGQPVIVRTAFNGSTPDLASAEEVAVAPDLQDTS 375
 Oy 374 OKQARLFLGIVAVDKQAPLPHLGRVHTAGMPAIVPTGERTVLYLFESPLDRLTETMR 433
 Db 376 RGTSTYVYLRIRPNKAGMKHLPGKILYRPMRAVEFIQTSERSVLYFYKPFDDRLKKTIV 435
 Oy 434 EE 435

Db 436 OE 437
 RESULT 12
 ID 054417 PRELIMINARY; PRT; 443 AA.
 AC 054417;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Metalloprotease transporter.
 GN HASE8000.
 OS Serratia marcescens.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Serratia.
 OX NCBI_TaxID=615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SR41;
 RA Omori K.;
 RT "Cloning of the hasDE8000 genes from Serratia marcescens Str41:
 secretion mechanism for the lipase and metalloprotease.";
 RL Submitted (FEB-1996) to the EMBL/Genbank/DBJ databases.
 DR EMBL; D83582; BAA12016.1; -
 DR InterPro: IPR002215; HLYD.
 DR InterPro: IPR003997; RtxD.
 DR Pfam: PF00529; HLYD; 1.
 DR PRINTS; PRO1450; RYTOXIND.
 DR PROSITE; PS00543; HLYD_FAMILY; 1.
 KW Protease.
 SQ SEQUENCE 443 AA; 48409 MW; 762FD7F822B0C44 CRC64;

Query Match 25.0%; Score 540.5; DB 2; Length 443;
 Best Local Similarity 31.7%; Pred. No. 4.3e-26;
 Matches 138; Conservative 85; Mismatches 178; Indels 35; Gaps 6;
 Oy 18 RIGYGIATLFGGLGMAFAPLDSAV-TANGVSAEVSODVOHLEGMALIKLVREGK 76
 Db 25 RIGGMVILVFGFGFLMGLAPLDKGVPSGVVAGNKKAVQHPBGSGVSOIQVHEGR 84
 Oy 77 VKAGVLFELDPTQANAAGITRNOYVALKAMEARLLAERDORPSISPPADLTSGORADP 136
 Db 85 VVAGVLFELDPTQANAAGITRNOYVALKAMEARLLAERDORPSISPPADLTSGORADP 144
 Oy 137 VARLADDEQAOTERRQTOGVDLMNORLOQOSIEBIDQOTGLKQOLGFIETDELD 196
 Db 145 VMSIMLLOQQLTSRRALQSELAIAESIAQSQALBGRISYASKOROKAMLOEOLGG 204
 Oy 197 LRLKLDKGVPRPRLLALE-----ARAGSLGSGRLTADRSKAVOGASPDQLKVRQI 249
 Db 205 MKRLAAGVYVANNRLDLEGOYAQIDGQASDPTGNLRL-----GRQLLEKLRAI 255
 Oy 250 --KQEFFEVSOSITETRVRLAEVTEKEEVASDAQRIKIVSPVNGTAQNLRFETGAVV 307
 Db 256 ORREYQKESVQSLEAVERKLDLDELNRLAKAABADLGHQYKASVAGTVVGLVFTGCVI 315
 Oy 308 RAERPLVLDAPDEAFVQAHFOPIDVDVNHMGVTEVRLPAFHSGAPDPERHDPVANA 367
 Db 316 GAGQQLMETVPSDRGLQVBARIPVELIDKVOGLVPELLEFSAFNSBTRVGEVYLVGA 375
 Oy 368 DRISDPKQARL-----LGIIVAVDKQAPLPHLGRVHTAGMPAIVPTGERTVLY 419
 Db 376 DRLTEKSGAPYYSRAKASEAGLQRLNLELRP-----GMPEGRITRGESMNY 427
 Oy 420 LFSPLDRLTETMR 435
 Db 428 LFKPLDRLTETMR 443
 RESULT 13
 ID 005198 PRELIMINARY; PRT; 473 AA.
 ID 005198

Query Match	24.88	Score 536.5	DB 2	Length 473
Best Local Similarity	28.58	Pred. No. 8.4e-26		
Matches 119, Conservative	96	Mismatches 199	Indels 3	Gaps 2

RESULT	14
0921J2	
ID	0921J2
AC	0921J2:
DT	01-DEC-2001 (TREMblrel. 19, Created)
DT	01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT	01-JUN-2002 (TREMblrel. 21, last annotation update)
DE	Alkaline protease secretion protein APRE.
GN	APRE OR RC0428.
OS	Rickettsia conorii.
OC	Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC	Rickettsiaceae; Rickettsiae; Rickettsia.
OX	NCBI_TaxID=781;
NN	[1]

Query Match	24.5%	Score	530.5	DB	16	Length	511
Best Local Similarity	30.0%	Pred	No. 2.2e-25				
Matches	135	Conservative	88	Mismatches	202	Indels	25
						Gaps	7

RESULT 15
Q9XB63
ID Q9XB63 PRELIMINARY; PRT; 443 AA.

DT 01-NOV-1999 (TrEMBLref. 12, last sequence update)
DT 01-JUN-2001 (TrEMBLref. 17, last annotation update)
DE PrtE protein.
GN PrtE.
OS *Erwinia amylovora*.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC *Erwinia*.
OX NCBI_TaxID=552;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E9;
RX MEDLINE=99303693; PubMed=10373365;
RA Zhang Y., Bak D.D., Heid H., Gelder K.;
RT "Molecular characterization of a protease secreted by *Erwinia*
RL *amylovora*.";
J. Mol. Biol. 289:1239-1251(1999).

```

DR  EMBL: Y19002; CAB42875.1; -.
DR  InterPro: IPR002215; HLVD.
DR  InterPro: IPR000847; HNH_LysR.
DR  InterPro: IPR003997; RXCD.
DR  Pfam: PF00529; HLVD.1
DR  PRINTS: PR01490; RRYTOXIND.
DR  PROSITE: PS00543; HLVD_FAMILY.1.
DR  PROSITE: PS00044; HNH_LYSR_FAMILY.
DR  SEQUENCE 443 AA; 48685 MW; 539AD285715994011 CRC64

```

Query Match	24.1%;	Score 522;	DB 2;	length 443;
Best Local Similarity	29.8%;	Pred. No. 6.3e-25;		
Matches 137; Conservative	81;	Mismatches 181;	Indels 60;	Gaps 9

[illegible]

Search completed: April 22, 2003, 12:32:22
Job time : 40 secs

THIS PAGE BLANK (per 10)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:25:06 ; Search time 25 Seconds

(without alignments)
721.688 Million cell updates/sec

Title: US-09-913-414-5

Perfect score: 2165

Sequence: 1 MKPPKIQRPDNDGAVARIG.....VLQYLFSPRLRLRTMKRE 435

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	514.5	23.8	432	1	APRE_PSEAE
2	506.5	23.4	448	1	PTE_ERWCH
3	313.5	14.5	477	1	RT3D_ACTPL
4	288.5	13.3	478	1	HLV4_ECOLI
5	267.5	12.4	477	1	HLVD_ACTAC
6	267.5	12.4	478	1	HLVD_ECOLI
7	266	12.3	440	1	CTAD_BORPE
8	243.5	11.2	478	1	HLVD_PASBP
9	233.5	10.8	477	1	RT1D_ACTPL
10	233	10.8	478	1	HLVD_PASHA
11	164	7.6	424	1	MCHE_ECOLI
12	163	7.5	413	1	CVAA_ECOLI
13	147.5	6.8	343	1	YTCR_ECOLI
14	146	6.7	421	1	EXOF_RHIME
15	142	6.6	355	1	YHII_ECOLI
16	136.5	6.3	382	1	YB94_HAEIN
17	130.5	6.0	879	1	MSP_ONCVO
18	130	6.0	1625	1	CMPI_MYCTU
19	129.5	6.0	848	1	MSP_DIRM
20	128	5.9	371	1	MACA_ECOLI
21	128	5.9	882	1	MSP_CAEEL
22	126.5	5.8	331	1	YBHG_SALTY
23	126	5.8	390	1	EMRA_ECOLI
24	125	5.8	331	1	YBHG_SALTY
25	124	5.7	371	1	MACA_YERPE
26	124	5.7	880	1	MSP_BRUMA
27	123.5	5.7	1576	1	POQC_AQUPE
28	122.5	5.6	378	1	YIBH_ECOLI
29	122	5.6	334	1	YJH3_RHIME
30	122	5.6	371	1	MACA_ECOS7
31	121.5	5.6	378	1	YIAY_ECOLI
32	121.5	5.6	545	1	HTR5_HALNI
33	121.5	5.6	4684	1	PLEI_HUMAN

ALIGNMENTS

RESULT 1	APRE_PSEAE	STANDARD:	PRT:	432 AA.
AC	003025:			
DT	01-OCT-1993 (Rel. 27, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Alkaline protease secretion protein apre.			
GN	APRE OR PA1247.			
OS	Pseudomonas aeruginosa.			
OC	Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;			
OX	NCBI_TaxID=287;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 15692 / PAOI;			
RX	MEDLINE=93051361; PubMed=1427098;			
RA	Duong F., Lazdunski A., Cami B., Murgier M.;			
RT	"Sequence of a cluster of genes controlling synthesis and secretion			
RT	of alkaline protease in Pseudomonas aeruginosa: relationships to			
RL	other secretory pathways."			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 15692 / PAOI;			
RX	MEDLINE=20437337; PubMed=10984043;			
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,			
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,			
RA	Barbey L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,			
RA	Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.;			
RA	Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;			
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an			
RT	opportunistic pathogen."			
RL	Nature 406:959-964(2000).			
RM	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 15692 / PAOI;			
RX	MEDLINE=20437337; PubMed=10984043;			
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,			
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,			
RA	Barbey L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,			
RA	Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.;			
RA	Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;			
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an			
RT	opportunistic pathogen."			
RL	Nature 406:959-964(2000).			
RM	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 15692 / PAOI;			
RX	MEDLINE=20437337; PubMed=10984043;			
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,			
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,			
RA	Barbey L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,			
RA	Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.;			
RA	Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;			
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an			
RT	opportunistic pathogen."			
RL	Nature 406:959-964(2000).			
RM	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 15692 / PAOI;			
RX	MEDLINE=20437337; PubMed=10984043;			
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,			
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,			
RA	Barbey L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,			
RA	Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.;			
RA	Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;			
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an			
RT	opportunistic pathogen."			
RL	Nature 406:959-964(2000).			
RM	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 15692 / PAOI;			
RX	MEDLINE=20437337; PubMed=10984043;			
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,			
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,			
RA	Barbey L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,			
RA	Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.;			
RA	Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;			
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an			
RT	opportunistic pathogen."			
RL	Nature 406:959-964(2000).			
RM	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 15692 / PAOI;			
RX	MEDLINE=20437337; PubMed=10984043;			
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,			
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,			
RA	Barbey L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,			
RA	Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.;			
RA	Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;			
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an			
RT	opportunistic pathogen."			
RL	Nature 406:959-964(2000).			
RM	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 15692 / PAOI;			
RX	MEDLINE=20437337; PubMed=10984043;			
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,			
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,			
RA	Barbey L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,			
RA	Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.;			
RA	Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;			
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an			
RT	opportunistic pathogen."			
RL	Nature 406:959-964(2000).			
RM	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 15692 / PAOI;			
RX	MEDLINE=20437337; PubMed=10984043;			
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,			
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,			
RA	Barbey L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,			
RA	Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.;			
RA	Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;			
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an			
RT	opportunistic pathogen."			
RL	Nature 406:959-964(2000).			
RM	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 15692 / PAOI;			
RX	MEDLINE=20437337; PubMed=10984043;			
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,			
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,			
RA	Barbey L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,			
RA	Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.;			
RA	Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;			
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an			
RT	opportunistic pathogen."			
RL	Nature 406:959-964(2000).			
RM	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 15692 / PAOI;			
RX	MEDLINE=20437337; PubMed=10984043;			
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,			
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,			
RA	Barbey L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,			
RA	Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.;			
RA	Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;			
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an			
RT	opportunistic pathogen."			
RL	Nature 406:959-964(2000).			
RM	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 15692 / PAOI;			
RX	MEDLINE=20437337; PubMed=10984043;			
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,			
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,			
RA	Barbey L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,			
RA	Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.;			
RA	Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;			
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an			
RT	opportunistic pathogen."			
RL	Nature 406:959-964(2000).			
RM	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 15692 / PAOI;			
RX	MEDLINE=20437337; PubMed=10984043;			
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,			
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,			
RA	Barbey L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,			
RA	Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.;			
RA	Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;			
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an			
RT	opportunistic pathogen."			
RL	Nature 406:959-964(2000).			
RM	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 15692 / PAOI;			
RX	MEDLINE=20437337; PubMed=10984043;			
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,			
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,			
RA	Barbey L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,			
RA	Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.;			
RA	Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;			
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an			
RT	opportunistic pathogen."			
RL	Nature 406:959-964(2000).			
RM	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 15692 / PAOI;			
RX	MEDLINE=20437337; PubMed=10984043;			
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,			
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,			
RA	Barbey L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,			
RA	Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.;			
RA	Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;			
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an			
RT	opportunistic pathogen."			
RL	Nature 406:959-964(2000).			
RM	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 15692 / PAOI;			
RX	MEDLINE=20437337; PubMed=10984043;			
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,			
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,			
RA	Barbey L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,			
RA	Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.;			
RA	Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;			
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an			
RT	opportunistic pathogen."			
RL	Nature 406:959-964(2000).			
RM	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 15692 / PAOI;			
RX	MEDLINE=20437337; PubMed=10984043;			
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,			
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,			
RA	Barbey L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,			
RA	Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.;			
RA	Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;			
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an			
RT	opportunistic pathogen."			
RL	Nature 406:959-964(2000).			
RM	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 15692 / PAOI;			
RX	MEDLINE=20437337; PubMed=10984043;			
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,			
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,			
RA	Barbey L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,			
RA	Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.;			
RA	Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;			
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an			
RT	opportunistic pathogen."			
RL	Nature 406:959-964(2000).			
RM	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 15692 / PAOI;			
RX	MEDLINE=20437337; PubMed=10984043;			
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,			
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,			
RA	Barbey L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,			
RA	Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.;			
RA	Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;			
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an			
RT	opportunistic pathogen."			
RL	Nature 406:959-964(2000).			
RM	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 15692 / PAOI;			
RX	MEDLINE=20437337; PubMed=10984043;			
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,			
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,			
RA	Barbey L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,			
RA	Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.;			
RA	Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;			
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an			
RT	opportunistic pathogen."			
RL	Nature 406:959-964(2000).			
RM	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 15692 / PAOI;			
RX	MEDLINE=20437337; PubMed=10984043;			
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,			
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,			
RA	Barbey L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,			
RA	Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.;			
RA	Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;			
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an			
RT	opportunistic pathogen."			
RL	Nature 406:959-964(2000).			
RM	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 15692 / PAOI;			
RX	MEDLINE=20437337; PubMed=10984043;			
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,			
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,			
RA	Barbey L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,			
RA	Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.;			
RA	Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;			
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an			
RT	opportunistic pathogen."			
RL	Nature 406:959-964(2000).			
RM	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 15692 / PAOI;			
RX	MEDLINE=20437337; PubMed=10984043;			
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,			
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,			
RA	Barbey L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,			
RA	Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.;			
RA	Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;			
RT	"Complete genome sequence of Pseudomonas aeruginosa PAOI, an			
RT	opportunistic pathogen."			
RL	Nature 406:959-964(2000).			
RM	[2]			
RP	SEQUENCE FROM N.A.			
RC				

FT DOMAIN 1 14 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 15 36 POTENTIAL.
 FT DOMAIN 37 432 PERIPLASMIC (POTENTIAL).
 FT CONFLICT 378 379 AL -> RV (IN REF. 1).
 SO SEQUENCE 432 AA; 48052 MM; 4C845C65DECF15EE CRC64;
 Query Match 23.8%; Score 514.5; DB 1; Length 432;
 Best Local Similarity 31.6%; Pred. No. 1.9e-25;
 Matches 137; Conservative 88; Mismatches 182; Indels 27; Gaps 11;
 QY 15 AAVARIGIATLVGLGMAFAFLDS-AVIANGVSAEVSODVQHLLEGMLKILVR 72
 DB 11 AAVARIGIATLVGLGMAFAFLDS-AVIANGVSAEVSODVQHLLEGMLKILVR 69
 QY 73 EGEKVKAGVLELDPQANAAAGITRNQYALKAMEARLLAEKQRPISFPADLTQSR 132
 DB 70 DGOHVEAGEPILIRPEPQARNVDSLNRVYANARLQARLAEYDGRRTLEGPAQLAQA 129
 QY 133 ADPMVAAIADQOAFERRQTIQGVYDLMAQRLQYQSEIEGIDRQTOG---LKQDLG 188
 DB 130 PLPLTIGERL-ELQRLHLSROTALANESALRANIEGRRAQLEGL-ROTEGNQRLQORL- 186
 QY 189 FIEDELIDRLKLYDKGLVPRRLALE-----ARAGSLSGISGRLTADRKAQVQASD 241
 DB 187 -LNSQLSGARDLAEQGYPRNQLLEQERQALAEVNARLSESSGREGQIR-----QSLAE 238
 QY 242 TOLKVRQIKQEFFEVSOSITETRVRLAEVTEKEVYASDAQRIKIVSPVNGTAQNLRF 301
 DB 239 AOMRIAQKEEYRKREYNGQLAETQVNAQTLEELSSAREYELRHAIEIRAPVSGYVAGLKV 298
 QY 302 TEGAVVRAEPLVDIAPEDEAFVIOAHQPTDVNVHGMKTEVLPFHSAGNDPBRH 361
 DB 299 TDGAVIGGELMTIVPRSDLEVEGQALVNLVDRHISGLVEMLFTFNSKIPRVTGE 358
 QY 362 DPAVAVADRIADPQOARLEIGIVRVYKQRLPHNG-RVTAGMPAQVIVPTGERTVQYL 420
 DB 359 VTWVSADRLDEONKQPYALRAQVDAAM-GKLKGLIDRGMAVQVEVFRGERSLNYL 417
 QY 421 FSPRLDRLRTMRE 434
 DB 418 FKPLPRAHVALAE 431
 RESULT 2
 PRTE_ERWCH STANDARD; PRT; 448 AA.
 AC P23597;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE Proteases secretion protein prte.
 OS Erwinia chrysanthemi.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Pectobacterium.
 OX NCBI_Taxid=556;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90228333; Pubmed=2184029;
 RA Letofte S., Delepelaitre P., Wandersman C.;
 RT "Protease secretion by Erwinia chrysanthemi: the specific secretion
 RT functions are analogous to those of Escherichia coli
 RT alpha-hemolysin".
 RL EMBL J. 9:1375-1382(1990).
 CC -1- FUNCTION: INVOLVED IN THE SECRETION OF PROTEASES A, B, C AND G.
 CC -1- SUBCELLULAR LOCATION: INNER MEMBRANE-BOUND.
 CC -1- SIMILARITY: BELONGS TO THE HLYD FAMILY OF SECRETION PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X53253; CA37343.1;
 DR EMBL; M60395; AAA63635.1;
 DR PIR; S12526; S12526.
 DR InterPro; IPR002215; HLYD.
 DR InterPro; IPR003997; RCD.
 DR Pfam; PF00529; HLYD.1.
 DR PRINTS; PR01490; RTXTQXND.
 DR PROSITE; PS00543; HLYD_FAMILY_1.
 KW Transmembrane; Inner membrane; Transport.
 FT DOMAIN 1 30 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 31 50 POTENTIAL.
 FT DOMAIN 51 448 PERIPLASMIC (POTENTIAL).
 SO SEQUENCE 448 AA; 49579 MM; 5112ACE1AFB84473 CRC64;
 Query Match 23.4%; Score 506.5; DB 1; Length 448;
 Best Local Similarity 28.8%; Pred. No. 6.2e-25;
 Matches 128; Conservative 93; Mismatches 191; Indels 33; Gaps 7;
 QY 8 PPTDMPQAVARIGYGIATLVGLGMAFAFLDSAVANG-VVSAEVSODVQHLLEGML 66
 DB 20 RASRDEERALLRGMVLVAGFGFLMALLAPLDKGVAVQGNVYSGNRKVIQHMGQIV 79
 QY 67 AKILVREGKVKAGVLELDPQANAAAGITRNQYALKAMEARLLAEKQRPISFPA 126
 DB 80 DRIVKDCDRAAGQVLLTTLNVAVARTTSEGISQYDOLIRAEARLLAEKQRPISFPA 139
 QY 127 DLTSQADPMVAAIADQOAFERRQTIQGVYDLMAQRLQYQSEIEGIDRQTOGKLDQ 186
 DB 140 RLQAKORPEMAIITAE-----DLRSQOSLKEIDIGVARSIDGETTS 185
 QY 187 LGFTE-----DELIDRLKLYDKGLVPRRLALEARAGSLSGISGRLTADR 232
 DB 186 LGALQKVMSSKQSEGAITLQDLGRPLAADNVPYPRNMLETERRLPAYSELQOTSEV 245
 QY 223 SKAVQASDTOLKVRQIKQEFFEVSOSITETRVRLAEVTEKEVYASDAQRIKIVSPV 292
 DB 246 GTRDRDIOQKRLIAQROOEYDEKVENSLSDVQKLNELVLSQREKADFNLANQVRAFA 305
 QY 293 GTAQNLREFTEGAVVRAEPLVDIAPEDEAFVIOAHQPTDVNVHGMKTEVLPFHS 352
 DB 306 GTVYDMKLTETGAVIARQGVMDVPEQDPLVGRIRPEVAVDWSGLPELFTASQ 365
 QY 353 AGNDPDERHDPVAVADRIADPQOARLEIGIVRVYKQRLPHNG-RVTAGMPAQVIVP 410
 DB 366 STTPRVPGTVLLSADRLVD-EKDGTPYGL-RQVSEKRSLSHGLEIKGMVQGVFR 423
 QY 411 TGERTVQYLFSPRLDRLRTMRE 435
 DB 424 TGRSFIVYLFKPLMDRMHIALTEE 448
 RESULT 3
 RT3D_ACTPL STANDARD; PRT; 477 AA.
 AC Q08633;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE RTX-III toxin determinant D (ApX-III) (Cytolysin IIID) (CLY-IIID)
 DE (Toxin RTX-III secretion protein D).
 GN APXIIID OR CLYIIID OR RTXID.
 OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Actinobacillus.
 OX NCBI_Taxid=715;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=Serotype 2; Pubmed=8494611;
 RX MEDLINE=93263992; Chang Y.-F., Shi J., Ma D.-P., Shin S.-J., Lein D.H.;


```

RT      "Molecular analysis of the Actinobacillus pleuropneumoniae RTX
RL      toxin-III gene cluster";
RN      DNA Cell Biol. 12:351-362(1993).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN-405 / Serotype 8;
RX      MEDLINE-95012630; PubMed-7927703;
RA      Jansen R., Bialek J., van Geel A.B.M., Kamp E.M., Gielkens A.L.J.,
RA      Smits M.A.;
RT      "Genetic map of the Actinobacillus pleuropneumoniae RTX-toxin (Apx)
RT      operons: characterization of the AprXIII operons.";
RL      Infect. Immun. 62:4411-4418(1994).
CC      -1- FUNCTION: INVOLVED IN THE TRANSPORT OF THE TOXIN RTX-III.
CC      -1- SUBCELLULAR LOCATION: Inner membrane-bound (POTENTIAL).
CC      -1- MISCELLANEOUS: THE SEQUENCE SHOWN IS THAT OF SEROTYPE 2.
CC      -1- SIMILARITY: BELONGS TO THE HLYD FAMILY OF SECRETION PROTEINS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
DR      EMBL; L12145; AAA21926.1; -
DR      EMBL; X80055; CAAS6360.1; -
DR      InterPro; IPR002215; HLYD.
DR      InterPro; IPR003997; RLYD.
DR      Pfam; PF00529; HLYD. 1.
DR      PRINTS; PRO1490; RTXTOXIND.
DR      PROSITE; PS00543; HLYD_FAMILY. 1.
KW      Cytolysis; Transport; Transmembrane; Inner membrane.
FT      DOMAIN 1 59
FT      TRANSMEM 60 77
FT      DOMAIN 78 477
FT      VARIANT 44 44
FT      VARIANT 132 132
FT      VARIANT 185 186
FT      VARIANT 376 376
SQ      SEQUENCE 477 AA; 54780 MW; D229F1CA7D7E90A CRC64;
Query Match 14.5%; Score 313.5; DB 1; Length 477;
Best Local Similarity 25.3%; Pred. No. 9e-13;
Matches 112; Conservative 88; Mismatches 190; Indels 53; Gaps 13.
QY      19 IGYGIATLTFGLGMAFAPLDASAVIANG-VSAEVSQDQVLEHGMGLAKILVREGKV 77
DB      60 IAYLMFLFLAIY-ISTISKEVYASATGKLVESHGSKREIKPIALYKDLFFVADGQFY 118
QY      78 KAGQVLEFL-----DPTQANAAGITRNQVVALKAMEARLARLARDORPSISF-PADLT 129
DB      119 EKGQLLMLTALGCDADAKOKTKVSLGLERLDQYRKSLYSI-EHNRPLDLDFNQADPD 176
QY      130 S-QRRDPVAARAIDE-----QAQFERRRTTIGQVDLMAQRLOYSLEIGDRL 178
DB      177 SVOEDKTAGARHLITDEQETWOKQYQKELAYQRRQAQKQVLANIRKYESASRIE--- 232
QY      179 QTQGLKDLGFIIEDELIDLRKLYDKGLVPRPLATLEAR---AGLSGSGISGRLATDRSK 234
DB      223 -----KELSLDKRLIYDKYSISKHELLAOENRKYEASSELVSQSHLEVESD 280
QY      235 AVQAGSDTQLKVRQTKQEFFEQVSSQISTETVRLAEVTEKEEVYASDAQRIK--ITVSPVN 292
DB      281 LKAKQEDLKVTLQFKSDILEKLOINQOREKQNTLELEKNE-----QRLASITRAFVS 334
QY      293 GTAQNLFEETEGCAVRAARPLVDIAPEDAEFVIAQNHQPTVDVNVNMGVTEVRLPAFVS 352
DB      335 GTVOOLKHTKSGGVVTTAETLMTAVIAPEDVDLEVSALIONKDVGEVETGOEAVIKETPPY 394
QY      353 AGNDPREHNDVAADRIQDPOKQARLEGLVARDVQRLPRHLRG-RYTAAGRAVQIYPT 411
DB      395 TRYGLYGVKVTITLDALENPQ-GLVFNYSIIEINKKTLTGDDKEIQIGSGSVIAETKT 453

```

```

OY 412 GERTVLYLFSPRLDRTLTMTRE 434
Db 454 GERSVTSFLSPLEESITESTLRE 476

RESULT 4
HLXA_ECOLI
ID HLXA_ECOLI STANDARD: PRT: 478 AA.
AC P09986;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hemolysin secretion protein D, chromosomal.
GN HLYD.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-J96 / Serotype O4;
RX MEDLINE=85234404; PubMed=3891743;
RA Fejmleee T., Pellett S., Welch R.A.;
RT "Nucleotide sequence of an Escherichia coli chromosomal hemolysin.";
RL J. Bacteriol. 163:94-105(1985).
RN [2]
RP TOPOLOGY.
RX MEDLINE=92357011; PubMed=1495479;
RA Schuelein R., Gentschev I., Mollenkopf H.-J., Goebel W.;
RT "A topological model for the haemolysin translocator protein HlyD.";
RL Mol. Gen. Genet. 234:155-163(1992).
CC -1- FUNCTION: INVOLVED IN THE TRANSPORT OF HEMOLYSIN A.
CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane.
CC -1- SIMILARITY: BELONGS TO THE HLYD FAMILY OF SECRETION PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
DR EMBL: M10133; AAA23977.1; -.
DR PIR: D24433; LEECD.
DR InterPro: IPRO002215; HLYD.
DR InterPro: IPR003997; RtxD.
DR Pfam: PF00529; HLYD; 1.
DR PRINTS: PR01490; RTXTOXIND.
DR PROSITE: PS00543; HLYD_FAMILY; 1.
KW Hemolysis; Transport; Transmembrane; CYTOPLASMIC (PROBABLE).
FT DOMAIN 1 59 SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN)
ET TRANSMEM 60 80
FT FT PERIPLASMIC (PROBABLE).
FT FT PERIPLASMIC (PROBABLE).
SQ SEQUENCE 478 AA; 54591 MW; 9E9EDC42ECB8089 CRC64;.

Query Match 13.3%; Score 288.5; DB 1; Length 478;
Best Local Similarity 25.28; Pred. No. 3,4e-11;
Matches 116; Conservative 94; Mismatches 192; Indels 59; Gaps 18;

OY 1 MKPPRTPPTDFQAVARIGYIIALTFV-GLLGMAFAAPLDASAVANGVS-AEVSODV 58
Db 49 IETPVSRRP-----RIKVAFFIMGFVIATILSLVG-----QVEIVATANGKLTLSSRSKEI 99

59 QHLEGGMALKLVREGCEVKVAGOVFEELDPQTQANAAGITRNQYVALKA---MEARLA 114
Db 100 KPIENSIYKKEIVKGESEVRKGVDVLKT--TALGAEADPTLKTOSSLDLARLEQINVYLIS 157

OY 115 ---ENDQRPSISFPADLTQS-RADPMVARAIADAEOAF-TERRQTIGQGVLD--MNAORL 167
Db 158 RSIEINTKLPEIKLPDEPYQNVSEBEVLRATLSLKREQSTWQNOKYQELNLDKKRARRL 217
```

QY 168 QYSEIESIDRQGLKQDLGFEDELIDRLKLYDKGLVPRRLALAEAR---AGSLSG 223
 Db 218 TILRLINFEVSVREKSRDL-----DFRSLHKQAIKHAHVEQENKYEANIELRV 270
 QY 224 SIGRLTADRKAQVGSASTOLKVRQIKQEFPEQVSQSTETERRVLAETEKEVVASDQK 283
 Db 271 YKSOEIQEIESLAKKEEYQVLTOLEFKNEIDKRLKQYTDSTELLTELEKNE---ERQ 326
 QY 284 RIKIVSPVNGTAQNLRFETEGAVVRAAEPLVDIAPEDEAFYIOAHFOPTDVDMVMGMV 343
 Db 327 ASVIRAPVSGKVKQQLKHTBEGVVTATETLAVIPEDEDTLEVTAIVQKKDIGFIWQMA 386
 QY 344 EVRLPAEHSAGNPDEPRHDPAVA-----DRISDPQKQARLFLGIVRVYKOLP---P 393
 Db 387 IIRKEAF-----PYTRYGLVCKVKNNINDAIED-OKLGLFENVIVSYVEENDLSTGK 438
 QY 394 HLGRVNTAGMPAQVIVPTGERTVQVLFSPRLDRLRTMRE 434
 Db 439 HI--PLSSGMVATAEIKTGMRSVISYLLSPLEESVTESLHE 477

RESULT 5

HLVD_ACTAC STANDARD: PRT: 477 AA.

AC P18790;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Leukotoxin secretion protein D.
 GN LKTD OR AALTD.
 OS Actinobacillus actinomycetemcomitans (Haemophilus actinomycetemcomitans).
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae; Actinobacillus.
 OX NCBI_TaxId=714;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-JF2;
 RX MEDLINE=90384843; PubMed=2402458;
 RA Guthmiller J.M., Craig E., Cagle M.P., Kolodrubetz D.;
 RT "Sequence of the lktD gene from Actinobacillus actinomycetemcomitans";
 RL Nucleic Acids Res. 18:5292-5292(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92072446; PubMed=1961107;
 RA Lally E.T., Golub E.E., Kieba I.R., Taichman N.S., Decker S., Berthold P., Gibson C.W., Demuth D.R., Rosenblum J.;
 RT "Structure and function of the B and D genes of the Actinobacillus actinomycetemcomitans leukotoxin complex";
 RL Microb. Pathog. 11:111-121(1991).
 CC -1- FUNCTION: INVOLVED IN THE TRANSPORT OF THE HEMOLYSIN/LEUKOTOXIN.
 CC -1- SUBCELLULAR LOCATION: Inner membrane-bound (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE HLVD FAMILY OF SECRETION PROTEINS.
 CC -1- SIMILARITY: THE N-TERMINAL (AA 1-153) IS SIMILAR TO E.PAECALIS PORE FORMING PROTEIN EBSA.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/announce/](http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch) or send an email to license@isb-sib.ch).
 CC
 DR EMBL: X53956; CA37907.1;
 DR PIR: S11315; S11315.
 DR PIR: B61378; B61378.
 DR InterPro: IPR002215; HLVD.
 DR InterPro: IPR003997; RLXD.
 DR Pfam: PF00529; HLVD.1.
 DR PRINTS: PR01490; RTXTOXIN.D.

DR PROSITE: PS00543; HLVD_FAMILY; 1
 KW Hemolysis; Cytolysis; Transport; Transmembrane; Inner membrane.
 FT DOMAIN 1 59 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 60 80 POTENTIAL.
 FT DOMAIN 81 477 PERIPLASMIC (POTENTIAL).
 SO SEQUENCE 477 AA; 54651 MW; 54F20128C4D6260 CRC64;

Query Match 12.4%; Score 267.5; DB 1; Length 477;
 Best Local Similarity 23.7%; Pred. No. 7e-10;
 Matches 108; Conservative 87; Mismatches 207; Indels 53; Gaps 15;

QY 6 IORPDMQAVARIGYGIATLPEVGLGMAFAPLDSAVINMGVVS-AEVSQDVQHEGG 64
 Db 49 IETPVSN--APRFVSYSLMFLTLAII-VSIFSNVEIATASGRFALSGRSKETPIENS 105
 QY 65 MLAKLIVEGKGVAGVGLFELDPQANAAGIRNOYVALKAEARL-----LAERDOR 119
 Db 106 LKHLIFAGEGVYKGGELLKLTALGAEADYLTKTSLSQAKLEEFKYSLSLEAVEKQDL 165
 QY 120 PSISE-PADLTSSORADPWAVARAIADEQAQETERRQTOGVYL--MNAORLYOSEIEGI 176
 Db 166 PILDFSKIDLRFMTENDQKRRTLLIEBQSTWQKORHQKTLLNKKKEAKSTYLARI--- 222
 QY 177 DRQTOGLKQDLGFEDE---LIDRLKLYDKGLVPRRLALAEARAGSLSGSIGRLTADRS 233
 Db 223 -----KKYEGLINTEGVRLDDERFALYKHAHKTVLDEE---NKYQDAINELEYKA 272
 QY 224 KAVQ-----GASDPTOLKVRQIKQEFPEQVSQSTETERRVLAETEKEVVASDQKRIK 286
 Db 273 SLMOYENVLAKKEQELVLTDFKNDILDKKQATDNVNLITFELDKN---NORQVSE 328
 QY 287 IVSPVNGTAQNLRFETEGAVVRAAEPLVDIAPEDEAFYIOAHFOPTDVDMVMGMVTEVR 346
 Db 329 IRAPVSGTVQDLKVTHTIDGVVTTAETLIMVVPEDSLEVTALONKDKGFKEQGEQEVVVK 388
 QY 347 LPAF-----HSAGNPDEPRHDPAVADRISDPQKQARLFLGIVRVYKOLPRLRG-RV 399
 Db 389 VEAPPYTRYGLVCKVKN-----ITLDAIEHP-KLGLFENVITTELDKRTLSTEKEIPL 441
 QY 400 TAGMPAQVIVPTGERTVQVLFSPRLDRLRTMRE 434
 Db 442 SAGMEITAEIKTGMRSVISYLLSPLEESIDKSLAE 476

RESULT 6

HLVD_ECOLI STANDARD: PRT: 478 AA.

AC P06739;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hemolysin secretion protein D, plasmid.
 GN HLVD.
 OS Escherichia coli.
 OC Plasmid Inc12 pHLV152.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.
 OX NCBI_TaxId=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC HESS J., WEIS W., VOGEL M., GOEBEL W.;
 RA HESS J., WEIS W., VOGEL M., GOEBEL W.;
 RT "Nucleotide sequence of a plasmid-encoded hemolysin determinant and its comparison with a corresponding chromosomal hemolysin sequence";
 RL FEBS Microbiol. Lett. 34:1-11(1986).
 RN [2]
 RP TOPOLOGY.
 RX MEDLINE=92357011; PubMed=1495479;
 RA Schuelein R., Gentschew I., Mollenkopf H.-J., Goebel W.;
 RT "A topological model for the haemolysin translocator protein HlyD";
 RL Mol. Gen. Genet. 234:155-163(1992).
 CC -1- FUNCTION: INVOLVED IN THE TRANSPORT OF HEMOLYSIN A.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane.
 CC -1- SIMILARITY: BELONGS TO THE HLVD FAMILY OF SECRETION PROTEINS.

```

OX NCBI_TaxID=520;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=18323;
RX MEDLINE=89091151; PubMed=2905265;
RA Glaser P., Sakamoto H., Bellau J., Ullmann A., Danchin A.;
RT "Secretion of cyclolysin, the calmodulin-sensitive adenylate cyclase-
RT haemolysin bifunctional protein of Bordetella pertussis.";
RL EMBL J.7:3997-4004(1988).
CC -!- FUNCTION: CYAD IS NECESSARY FOR TRANSPORT OF CALMODULIN-SENSITIVE
CC ADENYLATE CYCLASE-HEMOLYSIN (CYCLOLYSIN).
CC -!- SUBCELLULAR LOCATION: Inner membrane-bound (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE HLYD FAMILY OF SECRETION PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL: X14199; CAA32413.1; -.
CC DR PIR: S02387; BVBRCD.
CC DR InterPro: IPR002215; HLYD.
CC DR InterPro: IPR003997; RCD.
CC DR Pfam: PF00529; HLYD.1.
CC DR PRINTS: PR01490; RTXTOXIND.
CC DR PROSITE: PS00543; HLYD_FAMILY.1.
CC KW Hemolysis; Transport; Transmembrane; Inner membrane.
CC FT DOMAIN 1 55 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 56 75 PERIPLASMIC (POTENTIAL).
CC FT DOMAIN 76 440 POTENTIAL.
CC SO SEQUENCE 440 AA; 48011 MM; B786328A92D0D48 CRC64;
Query Match 12.3%; Score 266; DB 1; Length 440;
Best Local Similarity 26.2%; Pred. No. 7.9e-10;
Matches 110; Conservative 71; Mismatches 181; Indels 58; Gaps 16;

16 VARIGYIATLFWGLGMAAFAPLDSAVING-VVSAEVSODYOHLGGMILAKILVREG 74
57 LARI---LMALA-AGALVWSVVGKREIYVHAAGKVVPVGQSKIIAASGTGRVAVRLVADN 112
75 EKVKAGQVLFELDPTQANAAAGITRNQYVALKAMEARLLAEKRDQPSISPPADLTSGRAD 134
113 SRVAGDVILRLD-----AGVFAEERKKRVQAAQ--ARQDEARSAMIRALDITGRA- 162
135 PAVVAALADEQQAQTFERRKOTILOGVDLMAAQRLOQVQSIEGIDKQTGSKKQOLEFIEBL 194
163 PVLAELPADPGGMAAQ-----SYLDSQYADVQAOLRSIEAIAIATYRRDVGVLQVIA 213
195 IDLRKLYDGKGLVPRRLALAEARAGSLSGISGRLTADRSKAVOGASDPQLVVRKIQEFF 254
214 HAHRKLARDGVSSQAYILEKEDARKLTLEGR-----RQSEKORALALQVOTR-----KQAF- 263
255 EQVSOSITETRY--RLAEVTEKEVYVASDAQ-RIKIYSPVNGTAONLRFETEGAVYRAA 310
264 -----ETVLARKLAAQAEQELAPTSQAQRSLRVLPAADGVYQQLVALTEGFAVAVAT 315
311 EPLVIAPEDEAFVIOAFOPIDVDNVMHGNVTEKRLPAFHSAGNPDEPRHDPAVAVADRI 370
316 QPLMKVVSFGAGIQVQAOLDSKDIDGFVAGAPATKVGADYTRKXGTLEGVGLVYVSPDV 375
371 SDPOKO--ARLEFLG-----VRVDYQQLRPHLRGRVTAAGPQAVIYPTGERTVLOLFSPL 424
376 VDDRQGHSTYRTIALAHPALEYDGK--PRL-----LKESMAVQAODIRGSRRLIEYLLSPV 429

```

DT 01-OCT-1996 (Rel. 34, last sequence update)
 DT 15-JUN-2002 (Rel. 41, last annotation update)
 DE Leukotoxin secretion protein D.
 GN LKTD.
 OS Pasteurella haemolytica-like sp. (strain 5943b).
 CC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 CC Pasteurella.
 CC NCBI_TaxID=28165;
 RX MEDLINE=93239320; PubMed=8478098;
 RA Chang Y.-F., Ma D.-P., Shi J., Chengappa M.M.:
 RT "Molecular characterization of a leukotoxin gene from a Pasteurella
 RT haemolytica-like organism, encoding a new member of the RTX toxin
 RT family.";
 RL Infect. Immun. 61:2089-2095(1993).
 CC -1- FUNCTION: INVOLVED IN THE TRANSPORT OF THE LEUKOTOXIN.
 CC -1- SUBCELLULAR LOCATION: Inner membrane-bound (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE HLXD FAMILY OF SECRETION PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L12148; AAA16446.1;
 DR InterPro: IPR002215; HLXD.
 DR InterPro: IPR003997; RTX.
 DR Pfam: PF00529; HLXD.1
 DR PRINTS: PRO1490; RTXTOXIND.
 DR PROSITE: PS00543; HLXD_FAMILY.1
 KM Hemolysis: Cytolysis; Transport; Transmembrane; Inner membrane.
 FT DOMAIN 1 77 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 78 98 POTENTIAL.
 FT DOMAIN 99 PERIPLASMIC (POTENTIAL).
 SO SEQUENCE 478 AA; 54650 MW; 221461a69482913a CRC64;
 Query Match 11.2%; Score 243.5; DB 1; Length 478;
 Best Local Similarity 24.4%; Pred. No. 2.3e-08;
 Matches 113; Conservative 73; Mismatches 174; Indels 103; Gaps 18;
 QY 31 LLGNAAPPLDSAVIANCVSAEV-----SDDVHLEGMAKIVREGKV 77
 DB 59 LIALILMLFLAVLAVLSVSKVEIVATAPGKLTSGRSKEKPIENTIVOEIFVDCGFV 118
 QY 78 KAGOVLEFLDP--TOANAAGITRNOYVALKAMEARLL--AERDORSISFPDLT--- 129
 DB 119 EKGGLVSLTRAGSDADIKTTITSLAKLENRYQTLTAIEKESLPVI---DLSTTE 174
 QY 130 ---SORADPMVARAIADP---QAQFTER-----ROTTIOGVLDLNAAGRLQYOS 171
 DB 175 FKDSSEEDRLIKHLIEQYVTTWQKQKQKTLAVKRKAENQOTISA-----YVR 223
 QY 172 EIEGIDRQGLKQQLGFIEDLIDLRKLYDKGLVPRRLALAEARAGSLSGIGRLTAD 231
 DB 224 KYEGATRIEIO-----EKIKDFRLXYOKSKLSKHELLSOENKAIEAO---NELAVY 270
 QY 232 RSKAVOGASD-----TOLKVRQIKOEFEFVOVSITETFRVRLAEVKEVAVSDAOKR 284
 DB 271 RSKLNELESLDLNKEELLETTOFFKSDVLEKLOHLENEBOLEL-EKN---NQRQA 326
 QY 285 IKIVSPVNGTAQNLRFTEGAIVVRAAEFLVDIAEDFAVQAHFOPTDVNVHMGWTE 344
 DB 327 SMIRAPVSGYQQLKIHIGGVVTAETIMTIIVEDEDDLEFALIQNDIGFVAAQGVYI 386
 QY 345 VRLPAFHSAGNPDDERHDPAV---VADRIIS---DPQKQARL---FLGIVRVVYKQL- 391
 DB 387 IKVETFE-----PYTRGYITGRIRKIHISPADEQNLGLVFNAVATSIDKQALLS 434
 QY 392 PPHLRGAVTAGMPAOVIVPGERIVLOYLPSPLDITLTATRE 434

DB 435 PDGKHIEGLGMLTTAEIKTGERSVSYLSPLSESVESURE 477
 RESULT 9
 RTID.ACPTL STANDARD; PRT: 477 AA.
 ID RTID.ACPTL
 AC P26761;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, last sequence update)
 DT 15-JUN-2002 (Rel. 41, last annotation update)
 DE RTX-I toxin determinant D (APX-ID) (HLX-ID) (Cly-ID)
 DE (toxin RTX-I secretion protein D).
 GN APXID OR CLYID OR HLXD OR APPD.
 OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
 CC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 CC Actinobacillus.
 CC NCBI_TaxID=715;
 RX MEDLINE=91317735; PubMed=1860823;
 RA Chang Y.-F., Young R., Struck D.K.:
 RT "The Actinobacillus pleuropneumoniae hemolysin determinant: unlinked
 RT appca and appbd loci flanked by pseudogenes."
 RL J. Bacteriol. 173:5151-5158(1991).
 RN [12]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Serotype 5;
 RX MEDLINE=92040145; PubMed=1937809;
 RA Smiths M.A., Briatore J., Jansen R., Smith H.E., Kamp E.M.,
 RA Gielkens A.L.:
 RT "Cytolysins of Actinobacillus pleuropneumoniae serotype 9.";
 RL Infect. Immun. 59:4497-4504(1991).
 RN [13]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S 4074 / Serotype 1;
 RX MEDLINE=94237497; PubMed=8181764;
 RA Frey J., Haldimann A., Nicolet J., Boffini A., Prentki P.:
 RT "Sequence analysis and transcription of the apxi operon (hemolysin I)
 RT from Actinobacillus pleuropneumoniae.";
 RL Gene 142:97-102(1994).
 CC -1- FUNCTION: INVOLVED IN THE TRANSPORT OF THE TOXIN RTX-I AS WELL AS
 CC THAT OF RTX-II.
 CC -1- SUBCELLULAR LOCATION: Inner membrane-bound (POTENTIAL).
 CC -1- MISCELLANEOUS: THE SEQUENCE SHOWN IS THAT OF SEROTYPE 5.
 CC -1- SIMILARITY: BELONGS TO THE HLXD FAMILY OF SECRETION PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M65808; AAB00967.1;
 DR EMBL: X61112; CAA43426.1;
 DR EMBL: X68595; CAA48588.1;
 DR PIR: B40366; B40366.
 DR PIR: S18856; S18856.
 DR InterPro: IPR002215; HLXD.
 DR InterPro: IPR003997; RTX.
 DR Pfam: PF00529; HLXD.1.
 DR PRINTS: PRO1490; RTXTOXIND.
 DR PROSITE: PS00543; HLXD_FAMILY.1.
 KM Hemolysis: Cytolysis; Transport; Transmembrane; Inner membrane.
 FT DOMAIN 1 59 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 60 80 POTENTIAL.
 FT DOMAIN 81 477 PERIPLASMIC (POTENTIAL).
 FT VARIANT 14 14 A -> R (IN SEROTYPES 1 AND 9).
 FT VARIANT 94 94 G -> D (IN SEROTYPES 1 AND 9).
 FT VARIANT 113 113 E -> Q (IN SEROTYPES 1 AND 9).
 FT VARIANT 113 113

FT VARIANT 228 228 S -> I (IN SEROTYPES 1 AND 9).
 FT VARIANT 335 335 Y -> GT (IN SEROTYPES 1 AND 9).
 SQ SEQUENCE 477 AA; 54727 MW; B841D87AF7DAD4C2 CRC64;
 Query Match 10.8%; Score 233.5; DB 1; Length 477;
 Best Local Similarity 22.9%; Pred. No. 9.6e-08;
 Matches 105; Conservative 85; Mismatches 185; Indels 83; Gaps 17;
 QY 19 IGYGIATFVGLGMAFAPLDSAVIANGVVS-AEVSODVOHLEGGMLAKTIVEGK 77
 DB 60 IAYLMFLFALV-ISTVSHVEIVATGKIAFGSKRKPIENALVKELFVEDGQFV 118
 QY 78 KAGVLFELDPTQANAAGITRNQYVALKAMEAR--LLAE--RPDRSISPAD--LTS 130
 DB 119 EKDQLLHLTALGADADQOKSSLSLTKLERYRIEILLEVADRLLEITKDFEKNH 178
 QY 131 QRADPVARAIADEQAQFTER-----RQTIGGVDDLMNQRLQYSEIEGIDRQT 180
 DB 179 TEEDKTRIRYLITEQFEAMOKOKYQKELALORREAKOTVLANIRKYGSSHVE----- 232
 QY 181 QGLADQGLFIDELELDLRLKLDKGLVPRRLALEAR---AGLSGSGITRLADRKA 236
 DB 233 -----NRLKDLKLFNSKSTSHVDVLTQENRHIENAVNELAVYKSRINEVESDLR 282
 QY 237 QGASDTOLVKROIKEFEQVOSITETRVRLAEVTEKEVVASDAOKRIKIV--SPVNGT 294
 DB 283 QAKEIHLITOLFADILEKLKQNV-EAKQSLSELEK-----NQRQIASYTRAPVS-Y 335
 QY 295 AQNTFEFEGAVVRAAEPLVDIAPEDEAFVIOAHFOPTVDNVHMGVTEVRLPAFHSAG 354
 DB 336 VQQLKTHVGGVVTAETLVIATAPEDVDLEVLTALQNKDIGIEVGQAVIKVEF----- 391
 QY 355 NPDEPRHDPVA-----VADRI SDPQKARLFLGIVRVVYKQPLRLR----- 398
 DB 392 -----PYTRYGYLMKRVKNITLLEALEHPQ-LGLFVNSIISIDRKTL-----LSGDKGE 438
 QY 399 --VTAGMPAQVIVPTGERTVLOYLFSPRLDRLRTTMR 434
 DB 439 IELGSGMVTAEIKTGERSVISYLLSPLEESVSESLRE 476
 RESULT 10
 HLYD_PASHA STANDARD: PRT: 478 AA.
 AC P16534;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Leukotoxin secretion protein D.
 GN LKTD.
 OS Pasteurella haemolytica.
 CC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 CC Mannheimia.
 OC NCBI_TaxID=75985;
 OX 111
 RN RN
 RP SEQUENCE FROM N.A.
 RC STRAIN-Serotype A1;
 RX MEDLINE=891213172; PubMed=2914876;
 RA Strathdee C.A., Lo R.Y.C.;
 RT "Cloning, nucleotide sequence, and characterization of genes encoding
 RT the secretion function of the Pasteurella haemolytica leukotoxin
 RT determinant";
 RL J. Bacteriol. 171:916-928(1989).
 RN 131
 RP SEQUENCE FROM N.A.

RC STRAIN-Serotype A1 / PHL101;
 RA MEDLINE=89210283; PubMed=2707120;
 RX Highlander S.K., Chidambaram M., Engler M.J., Weinstock G.M.;
 RT "DNA sequence of the Pasteurella haemolytica leukotoxin gene
 RT cluster";
 RL DNA 8:15-28(1989).
 CC -1- FUNCTION: INVOLVED IN THE TRANSPORT OF THE LEUKOTOXIN.
 CC -1- SUBCELLULAR LOCATION: Inner membrane-bound (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE HLYD FAMILY OF SECRETION PROTEINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@isb-sdb.ch).
 CC
 DR EMBL: M20730; AAA25531.1;
 DR EMBL: M24197; AAA25545.1;
 DR PIR: B32051; B32051.
 DR PIR: S29518; S29518.
 DR InterPro: IPR002215; HLYD.
 DR InterPro: IPR003997; RtxD.
 DR Pfam: PF00529; HLYD; 1.
 DR PRINTS: PR01490; RYTXOIND.
 DR PROSITE: PS00543; HLYD_FAMILY_1.
 KW Hemolysis; Cytolysis; Transport; Transmembrane; Inner membrane.
 FT DOMAIN 1
 FT TRANSMEM 60 80
 FT POTENTIAL. CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 81 478
 FT PERIPLASMIC (POTENTIAL).
 FT CONFLICT 18 18 T -> I (IN REF. 3).
 FT CONFLICT 47 47 D -> E (IN REF. 3).
 FT CONFLICT 150 150 N -> T (IN REF. 3).
 FT CONFLICT 236 236 F -> L (IN REF. 3).
 FT CONFLICT 267 267 EL -> AV (IN REF. 3).
 SQ SEQUENCE 478 AA; 54761 MW; E2B8B5FF5AF988C1 CRC64;
 Query Match 10.8%; Score 233; DB 1; Length 478;
 Best Local Similarity 23.7%; Pred. No. 1e-07;
 Matches 109; Conservative 92; Mismatches 174; Indels 84; Gaps 20;
 QY 19 IGYGIATFVGLGMAFAPLDSAVIANGVVS-AEVSODVOHLEGGMLAKTIVEGK 77
 DB 60 IAY-LIMFLVVAIVLASKEYEIVATAPGKLTGSGRSKEIPENALVQELFVADGQFV 118
 QY 78 KAGVLFELDPTQANAAGITRNQ-YVALKAME---ARLLA--BRDQPSISPADLT- 129
 DB 119 EKDQLLVSL--TALGSDADIKKTMASTSLAKLENYRYQTLLTAIKESLPVTL--DLR 172
 QY 130 -----SQRADPVARAIADEQAQFTERQTTQGVDDLMNAOR-----LQYSEIEGID 177
 DB 173 TEFKDSSEDRIRIKHLIEQYTTWQOKT---QKTLVKKREAKQTIFAVRYREGAT 229
 QY 178 ROTGIGKQDGLFIDELELDLRLKLDKGLVPRRLALEARAGLSGSGITRLADRSKAVQ 237
 DB 230 RLEQ-----EKPRFKALVYKQKSKSKHELLAQE---NKLLEAQNELAVYSKLNE 276
 QY 238 GASD-----TQLKRVROIKEFEQVOSI--TETRVRLAEVTEKEVVASDAOKRIKIV- 288
 DB 277 LENDLLNVKEELLEITQFFKSDVLEKLKQHTENRQLRL-----ELEKNNQRQASMR 330
 QY 289 SPVNGTAQNLRFETEGAVVRAAEPLVDIAPEDEAFVIOAHFOPTVDNVHMGVTEVRLP 348
 DB 331 APVSGTVQQLKHTIGVAVTAETLMTIIPEDVDLEATLVPNKDIGFAVAAQOEVIKVE 390
 QY 349 AF-----HSAG-----NPDEPRHDPVAVADRI SDPQKARLFLGIVRVYKQPL- 395
 DB 391 TFPYTRYGYLGRIRKIHISPDALQEPNGL-----VFNAITAIADKRLNLSPOGR 438
 QY 396 RGRVTAGMPAQVIVPTGERTVLOYLFSPRLDRLRTTMR 434
 DB 439 KIDLSGMTVAETAEIKTGERSVISYLLSPLEESVSESLRE 477

```

RESULT 11
MCH_ECOLI
ID MCH_ECOLI STANDARD; PRT: 424 AA.
AC 09EXN6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Microcin H47 secretion protein mchE.
GN MCH_E.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-H47;
RX MEDLINE=21091907; PubMed=11181394;
RA Apptroz M.F., Rodriguez E., Lavina M.;
RT "The structure, function, and origin of the microcin H47 ATP-binding cassette exporter indicate its relatedness to that of colicin V."
RL Antimicrob. Agents Chemother. 45:969-972(2001).
CC -1- FUNCTION: PROBABLY INVOLVED, IN CONJUNCTION WITH MCH_F, IN THE SECRETION OF MICROICIN H47.
CC -1- SUBCELLULAR LOCATION: INNER MEMBRANE-BOUND (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE HLYD FAMILY OF SECRETION PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ278866; CAC21493.1;
DR InterPro: IPR002215; HLYD.
DR Pfam: PF00529; HLYD; 1.
DR PROSITE: PS00543; HLYD_FAMILY; 1.
KW Transport; Protein transport; Bacteriocin transport; Transmembrane;
KW Inner membrane.
FT DOMAIN 1 25 CYTOPLASMIC (POTENTIAL).
FT TRANSEM 26 46 POTENTIAL.
FT DOMAIN 47 424 PERIPLASMIC (POTENTIAL).
SQ SEQUENCE 424 AA; 48719 MW; 3CADD07D9EC3A41P4 CRC64;

Query Match
Best Local Similarity 21.7%; Pred. No. 0.0019;
Matches 93; Conservative 73; Mismatches 142; Indels 120; Gaps 18;

QY 50 VSAEVS---ODVOHLEG--GMLAKILVREGKRVAGQVLEFELDPTQANAAGITRNOYVA 104
DB 55 VSGEVTWTPRAVNINYSVGQVGFVQAFVHEGLIKKGPVYIDISKRSRISIVDHN--- 111
QY 105 LKAMEALLAERDRSPISPPADITSO--RADPVARAIDECQAQETERRTOTIGQVDLM 162
DB 112 -----RRDIENQIVRVNITISR--LEESKIT-----DITL 140
QY 163 NAORLOY-----QSELEGIDROTGKLDQGLFEDELIDIRKLYDGLVPRRL---LA 213
DB 141 EKORLQQTDFRNSSDI--IQRAEGIK---IKNNMNNRNYQAGLKKDQLNQOVA 194
QY 214 L---EARRAGSLG---SIGRLTADRSKAVOGASDTOLKVKROIKOFFEVSQSITETR 265
DB 195 LYYQQQNNLSTSGQNNQMLQITLESQIOTQADPDNRIRYQELORFE----- 244
QY 266 VRLAEVTEKEVVASDAQRIKIVSPVNGTAQNLFFTEGAVRAAEPLVDIAPDEAFVI 325
DB 245 -----LQKELVNTDVEGEIIRALTDGKVDLSL-VTVGQVNVGDSLLQVTPENIENY 297
QY 326 QAHPQPDV-DNVNMGAVTEYRLPAFHS-----AGNDP 358

```

```

DB 298 LILWPNDAVPIYISAGDKVNIIRYAPFAKFGFSATVYKTSRKPASTQEMLYKAPQN 357
QY 359 ERHDPVAVADRISDPQKARFLGIVRVYKOLPPLHGRVYTAGMPQAVIPGERTVLO 418
DB 358 TPGASVPMYKVIAMPKQ-----IIRDEKYL-----LENMKASSTLEKRRRIYQ 405
QY 419 YLFSPLRD 426
DB 406 WMLSPFYD 413

RESULT 12
CVAA_ECOLI
ID CVAA_ECOLI STANDARD; PRT: 413 AA.
AC P22519;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Colicin V secretion protein cvaa.
GN CVAA.
OS Escherichia coli.
OC Plasmid Incret ColV3-K30.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=91065315; PubMed=2249654;
RA Gilson L., Mahanty H.K., Koller R.;
RT "Genetic analysis of an MDR-like export system: the secretion of colicin V."
RL EMO J. 9:3875-3884(1990).
RN 12
RP TOPOLOGY.
RX MEDLINE=96042091; PubMed=7592380;
RA Skvitsky R.C., Reginald S., Shen X.;
RT "Topology analysis of the colicin V export protein cvaa in Escherichia coli."
RL J. Bacteriol. 177:6153-6159(1995).
CC -1- FUNCTION: INVOLVED, IN CONJUNCTION WITH CVAB, IN THE SECRETION OF COLICIN V.
CC -1- SUBCELLULAR LOCATION: Inner membrane-bound (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE HLYD FAMILY OF SECRETION PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X57524; CAA0743.1;
DR PIR: S12271; IRECSA.
DR InterPro: IPR002215; HLYD.
DR Pfam: PF00529; HLYD; 1.
DR PROSITE: PS00543; HLYD_FAMILY; 1.
KW Plasmid; Transport; Protein transport; Bacteriocin transport; Transmembrane; Inner membrane; Alternative initiation.
KW Transmembrane; Inner membrane; COLICIN V SECRETION PROTEIN CVAA.
FT CHAIN 1 413 COLICIN V SECRETION PROTEIN CVAA*.
FT INIT_MET 161 161 FOR CVAA*.
FT DOMAIN 1 20 CYTOPLASMIC (PROBABLE).
FT TRANSEM 21 36 PROBABLE.
FT DOMAIN 37 413 PERIPLASMIC (PROBABLE).
SQ SEQUENCE 413 AA; 47344 MW; E22B855F0C4D951 CRC64;

Query Match
Best Local Similarity 22.4%; Pred. No. 0.0022;
Matches 96; Conservative 73; Mismatches 139; Indels 120; Gaps 20;

QY 50 VSAEVS---ODVOHLEG--GMLAKILVREGKRVAGQVLEFELDPTQANAAGITRNOYVA 104

```

```

Db 44 VSGEVTTPRAVNIYSGVGFVRFVHEGOLIKKDPYLLIDSKS-----TRN----- 93
QY 105 LKAMEARLLAERDORPSISFPADLTSQ--RADPMVARATADQOAFTERQTIQGVYDLM 162
Db 94 -----GVTYDNHRR-----DIENQLVKNITSR--LEESKIT-----LDLTL 129
QY 163 NNAORLYQ-----OSEIEGIDRQFOGLKDQOLFIEDELIDRLKLYDKGLVPPRL---LA 213
Db 130 EKQRIQYIDAFRRSSDI--IQRAEGIK-----IMKNMNMNYVYOSKGLINKDQLNVOYA 183
QY 214 L-----EARGSLSG-----SIGRLTADRSKAVGASDTOLKVRQIKOEFEYVSOSITERR 265
Db 184 IYQOQNNLISGONEQNALQITLESQIQOAFDFRITRYOMELQRL----- 233
QY 266 VALAEVEKEVVAQAQRIRKIVSPVNGAQLRFTTEGAVVRAAEPLVDIAPDEAFYI 325
Db 234 -----LQKELVNTVEGEIIRALSDGKVDLSL--VTQGMVNTGDSLIQVLEPENNY 286
QY 326 QAHPQPTD-VDNVHMGVTEVRLPAFHS-----AGNPDP 358
Db 287 LILWMPNDVAPYISAGDKVNIRYEAFPSSEKFGFSATVKTISRPASTQEMLTLYGAPQN 346
QY 339 EKHDPVNAADRISDPQKARLFLGIVRDVQOLPPLHGRVTAGMPAQVYVPTGERTVQ 418
Db 347 TPGASVPMYKVIATEPEKQ-----IIRYDEKYL-----LENGMKAESTLFLKRRRIYQ 394
QY 419 YLESPFLD 426
Db 395 WMLSPFYD 402

```

RESULT 13

YJCR_ECOLI

ID YJCR_ECOLI STANDARD: PRT: 343 AA.

```

AC P32716:
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein yjcr.
GN yjcr OR B4082.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=94089392; PubMed=8263357;
RA Blattner F.R., Burland V.D., Plunkett G. III, Sofia H.J.,
RA Daniels D.L.;
RT "Analysis of the Escherichia coli genome. IV. DNA sequence of the
RT region from 89.2 to 92.8 minutes."
RL Nucleic Acids Res. 21:5408-5417(1993).
CC -1- SUBCELLULAR LOCATION: Inner membrane-bound (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE HLYD FAMILY OF SECRETION PROTEINS.
CC -1- This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: U00006: AAC43176.1:
DR EMBL: AE000481: AAD13465.1:
DR EcoGene: E81954: yjcr.
DR InterPro: IPR002215: HlyD.
DR Pfam: PF00529: HlyD.
DR TIGFAMS: TIGR00998: Baa0101.1.
KW Hypothetical protein; Transmembrane; Inner membrane; Transport;
KW Complete proteome.
FT DOMAIN 1 12 CYTOPLASMIC (POTENTIAL).

```

```

FT TRANSMEM 13 33 POTENTIAL.
FT DOMAIN 34 343 PERIPLASMIC (POTENTIAL).
SQ SEQUENCE 343 AA: 36908 MW: C59988707C475FFE CXC64.

```

Query Match: 6.8%; Score 147.5; DB 1; Length 343;
 Best Local Similarity 22.5%; Pred. No. 0.016;
 Matches 78; Conservative 53; Mismatches 131; Indels 85; Gaps 14;

```

QY 4 PKIQRTDNFQAVARIGVGIITLTFYGLLGMAFAPLDSAVANGVSNESVDYQHLEG 63
Db 5 PK-KAPRSKFPALLVALVALVAV--TWR---VDSAPSTNAYASADTIDVAPVYS 56
QY 64 GMLAKLIVREGKAVAGVLFELDPT--QANAAGITRQVYALKAMEARLLAERQORPS 121
Db 57 GRIVELAVTDNVAQVGGDLFRIDPRPYANLA-----KAESLAIDQIMLTQRSVA 111
QY 122 ISFPADLTSQADPMVARAIADEQAQFERRQTIQGVYDLMNAORLYOSEIEGIDRQ 181
Db 112 QQFGAD--SVNATVEKARAA--KQATDTLR-----TE 141
QY 182 GLKDQGLFIEDLIDRLKLYDKGLVPRRLLLLEAR--AGSLSGSIGRLTADRSKAVQGS 240
Db 142 PLTKE-GEVASADVDPAKTAQRAEADLNAVLQASASAVSGDALVAQRA----- 193
QY 241 DTQLKVRQIKOEFEYVSOSITETRYLAETVEKEVVASDAQRIKIYSPVNGTAQNLRF 300
Db 194 -----AVENDIALTLHLEMA-----VRAPFDGRVYISLK- 223
QY 301 FTGAVVRAAEPLVDIAPDEAFVIOAHFQPPDVNDVHMGVTEVRL 347
Db 224 TSVGQFASAMRPIFTLIDTRHMYVI-ANREFTDLKNIRSGTATIRL 269

```

RESULT 14

EXOF_RHIME

ID EXOF_RHIME STANDARD: PRT: 421 AA.

```

AC Q02728:
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Exopolysaccharide production protein exoC precursor.
GN EXOF OR RB1068 OR SMB20945.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RCR2011 / SU47;
RX MEDLINE=93177026; PubMed=8439670;
RA Mueller P., Keller M., Weng W.M., Quandt J., Arnold W., Puehler A.;
RT "Genetic analysis of the Rhizobium meliloti exoC operon: ExoC is
RT homologous to sugar transferases and ExoD represents a transmembrane
RT protein."
RL Mol. Plant Microbe Interact. 6:55-65(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396508; PubMed=11481431;
RA Finan T.M., Weidner S., Wong K., Buhmester J., Chain P.,
RA Vorholter F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
RA Golding B., Puehler A.;
RT "The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-
RT fixing endosymbiont Sinorhizobium meliloti."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
CC -1- FUNCTION: INVOLVED IN SUCCINOGLYCAN (EPS I) SYNTHESIS. NEEDED FOR
CC THE ADDITION OF THE FIRST SUGAR (GALACTOSE) TO THE ISOPRENOID
CC CARRIER.
CC -1- PATHWAY: Exopolysaccharide biosynthesis.
CC -1- SUBCELLULAR LOCATION: Periplasmic (probable).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L05588; AAA26265.1; -
 DR EMBL: AL003645; CAC949468.1; -
 KW Exopolysaccharide synthetase; Plasmid; Periplasmic; Signal;
 FT SIGNAL 1 31 POTENTIAL.
 FT CHAIN 32 421 EXOPOLYSACCHARIDE PRODUCTION PROTEIN
 FT SEQUENCE 421 AA; 45871 MW; 55D8EC55B3A469EF CMC64;
 SO
 Query Match 6.7%; Score 146; DB 1; Length 421;
 Best Local Similarity 22.3%; Pred. No. 0.026;
 Matches 66; Conservative 47; Mismatches 121; Indels 62; Gaps 6;
 QY 22 GIIATFTGILGMAFAFLDSAVINGVSAEVSODVHLEBGMIAKI----- 69
 Db 79 GSLSLPEFGDIP-----ASGKTSEVAEEI---GVKQKLFGLRDRPSASVE 122
 QY 70 -----LVREGKVKAGQ-----VL-----FELDPTQANNAAGIT 98
 Db 123 MAQRPVYLSEVOTPGETPYAPNLTLYKANVSLGGGLRADNGOFARDYINASESAVQ 182
 QY 99 RNQYVALKAMEARLLAEHDQRPISFPADLTSSRADPVAARADDEQAQFERRQTOGQ 158
 Db 183 VAERSRLIRARRLLAEIGKRDITPMPELKNV---PDAEKLDSEFALMESRDKRQKQ 239
 QY 159 VDLNMAQLOQOSEIEGIDRTQGLKDLGFIEDLIDRLKLYDKGLVPRRLALEKRA 218
 Db 240 LDALADLSELLQSEIEALAKKAETQARQLATELDRDQVDSIAEKGLALSORKLSEQRV 299
 QY 219 GSLSGSIKRLTADRSKAVQASDTQKVKROIKEFEFQVOSITETRVRLAEVPEK 274
 Db 300 ADVQASLIDITFASLKAKQDASKAAQDETNRNDMDQALQELQTEALDTLTILK 355
 RESULT 15
 ID YHII_ECOLI STANDARD; PRT; 355 AA.
 AC P37626;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein yhiI precursor.
 GN YHII OR B3487.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE-94316500; Pubmed-8041620;
 RA Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.;
 RT "Analysis of the Escherichia coli genome. V. DNA sequence of the
 RT region from 76.0 to 81.5 minutes."
 RL Nucleic Acids Res. 22:2576-2586(1994).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U00039; AAB18463.1; -
 DR EMBL: AF000424; AAC76512.1; -

DR EcoGene; EG12224; yhiI.
 DR InterPro; IPR002215; HlyD.
 DR Pfam; PF00529; HlyD; 1.
 KW Hypothetical protein; Signal; Complete proteome.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 355 HYPOTHETICAL PROTEIN YHII.
 FT SEQUENCE 355 AA; 38828 MW; 6784DF1F19D5F5FA CMC64;
 SO
 Query Match 6.6%; Score 142; DB 1; Length 355;
 Best Local Similarity 21.2%; Pred. No. 0.037;
 Matches 87; Conservative 65; Mismatches 158; Indels 100; Gaps 20;
 QY 22 GIIATFTGILGMAFAFLDSDAVINGVSAEVSODVHLEBGMIAKILVREGKVK 78
 Db 14 GILAV--AAIVAMWLLRPAGVPEGFVSNGRIEA-TEVDINASKIAGRITDITLVKEGFVR 70
 QY 79 AGQVLELDPTQANNAAGITRNQYVALKAMEARLLAEHDQRPISFPADLTSSRADPWA 138
 Db 71 EGEVLAKMD-----TR-----VLQEQRLMAIQKEAQ----- 99
 QY 139 RAIDDEQAQFERRQTOGQVDLNNQRLQYQSEIEGI-DKQTOGLK-DQLGFIEDLID 196
 Db 100 -AVAAQALILEORQSETRAAQSLVN---QRAELDSYAKKRHTSRSLAQGALSAQQLD 154
 QY 197 LRKLVDKLVPRRLALEARAGSLSGSLGRLTADRSKAVQASDTQKVKROIKEFEQ 256
 Db 155 -----DRAAASARAALESASAKAQSASKAIEARNTIIQ----- 190
 QY 257 VSQSTETTRVRLAEVTEKEVVAASDAQRIKIVSPVNGTAQMLRFELEGAVVRAAEPLVDI 316
 Db 191 -----AQTRVBAQAQTERIAADIDDSLEK--APRDSGVQ-YRAAPEVLAAGRVLM 242
 QY 317 APEDEAFYQAHFOPTD-VDNVHMG-----MTEVRLPA---FHSAGNPDEPRHD 363
 Db 243 V--DLSVYMTFFLEPTEQAGTLKLGGEARLLIDAAPDIRIPATISFVASVAQFTPK---T 297
 QY 364 VAVADRISDPQKQARLFLGIYRVVQKLPRLRGVGTGMAQVIYPTGE 413
 Db 298 VETSD-----ERLKMFRVKARIPPELLQHL-EYKGLDGVAMVRVNE 341

Search completed: April 22, 2003, 12:31:38
 Job time : 28 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2003, 12:29:17 ; Search time 21 Seconds
(without alignments)
1991.356 Million cell updates/sec

Title: US-09-913-414-5
Perfect score: 2165
Sequence: 1 MKPPKIQRPDMFQAVARIG.....VLQYLFSPLRDLTFTMREE 435

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR_73:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1984.5	91.7	436	2	B87374
2	589	27.2	474	2	P86370
3	584	27.0	437	2	A82972
4	583	26.9	439	2	C96003
5	568	26.2	434	2	A83099
6	568	26.2	436	2	G98187
7	564	26.1	443	2	B83221
8	543.5	25.1	473	2	C95953
9	542.5	25.1	437	2	H97683
10	542.5	25.1	437	2	A12908
11	530.5	24.0	511	2	D97753
12	519.5	24.0	511	2	D71687
13	514.5	23.8	432	2	G26697
14	514.5	23.8	432	2	G83489
15	506.5	23.4	448	2	S12526
16	476.5	22.0	442	2	A50477
17	471.5	21.8	437	2	B49933
18	438.5	20.3	452	2	A82965
19	438.5	20.3	497	2	G98318
20	368.5	17.0	387	2	AG0835
21	368	17.0	481	2	D82381
22	367	17.0	395	2	D83412
23	329	15.2	475	2	B81047
24	325	15.0	391	2	D85547
25	325	15.0	391	2	H90696
26	314	14.5	473	2	P82561
27	308.5	14.2	478	1	SA8045
28	288.5	13.3	478	1	LEBCD
29	288	13.3	507	2	AD2140

30	267.5	12.4	477	2	B61378
31	267.5	12.4	478	2	S10058
32	266	12.3	440	1	BYBRCD
33	264	12.2	438	2	T03513
34	262	12.1	478	2	E43599
35	261.5	12.1	454	2	G81828
36	260	12.0	500	2	A82517
37	251.5	11.6	479	2	T00229
38	251	11.6	467	2	G82198
39	248	11.5	498	2	AH2468
40	234	10.8	478	2	D30169
41	215.5	10.0	583	2	S75805
42	203.5	9.4	425	2	AB1018
43	201.5	9.3	425	2	T14873
44	196	9.1	455	2	D95923
45	194	9.0	512	2	AB2047

ALIGNMENTS

RESULT 1

B87374

RsaA secretion system, membrane protein RsaA [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C:Accession: B87374

R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: B87374

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-436 <STO>

A:Cross-references: GB:AE005673; NID:q13422299; PIDN:AKK22993.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC1009

Query Match	91.7%	Score 1984.5	DB 2	Length 436
Best Local Similarity	93.6%	Pred. No. 7.9e-116		
Matches 408	Conservative 3	Mismatches 24	Indels 1	Gaps 1
QY	1	MPKPKIQRTDMFQAVARIGYGLTFTFGYLGMAAFPLDSAVIANGVSAEVS-QDVQ	59	
DB	1	MPKPKIQRTDMFQAVARIGYGLTFTFGYLGMAAFPLDSAVIANGVSAEVS-QDVQ	60	
QY	60	HEGGLAKILVREGEKVAGVLEFELDPQANAAAGITRNQYVALKAEARLLAERDQ	119	
DB	61	HEGGLAKILVREGEKVAGVLEFELDPQANAAAGITRNQYVALKAEARLLAERDQ	120	
QY	120	PSISPPADITSGRADPMVARAIADEQAQFERRQITQGVDDLMNQRLQYSEIEIDRQ	179	
DB	121	PSISPPADITSGRADPMVARAIADEQAQFERRQITQGVDDLMNQRLQYSEIEIDRQ	180	
QY	180	TGGLDKQGFIDEDELTKLTKYDKGLVPRRLALEREQASLSGSLTARSAVQGA	239	
DB	181	TGGLDKQGFIDEDELTKLTKYDKGLVPRRLALEREQASLSGSLTARSAVQGA	240	
QY	240	SPTQLKVRQIKQFEFEQVQSITETFRVRLAEVTEKEVVASDAQRIKIVSPVNGTQNL	299	
DB	241	SPTQLKVRQIKQFEFEQVQSITETFRVRLAEVTEKEVVASDAQRIKIVSPVNGTQNL	300	
QY	300	FTEGAIVRAAPLVDIADEDAFYIAHFQPTDVNVHMGAVTEVRLPAFHSAGNPDE	359	
DB	301	FTEGAIVRAAPLVDIADEDAFYIAHFQPTDVNVHMGAVTEVRLPAFHSAGNPDE	360	
QY	360	RHDVAVADRIQPOKQALFLGIVVDVKOLPPLHGRGVTGMPQOVVPGERTVLQY	419	
DB	361	GTIQSLSDRIQPOKQALFLGIVVDVKOLPPLHGRGVTGMPQOVVPGERTVLQY	420	

QY 420 LFSPLRDLRTTMRRE 435
 Db 421 LFSPLRDLRTTMRRE 436

RESULT 2

F98310

rhizobium secretion protein rspe (AF141932) [imported] - Agrobacterium tumefaciens (st
 C:Species: Agrobacterium tumefaciens
 C:Date: 22-Oct-2001 #sequence.revision 22-Oct-2001 #text.change 11-Jan-2002

C:Accession: F98310
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughly, D.; Scott, C.; Lappas, C.; Markelz, B.;

Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
 A:Reference number: A97359; PMID:11743194

C:Accession: F98310
 A:Status: preliminary
 A:Molecule type: DNA

A:Residues: 1-474 <KUR>
 A:Cross-references: GB:AE007870; PIDN:AAK90008.1; PID:915159979; GSPDB:GN00170

C:Genetics:
 A:Gene: AGR_L_2881

A:Map position: linear chromosome

Query Match 27.2%; Score 589; DB 2; Length 474;
 Best Local Similarity 31.9%; Pred. No. 3.2e-29;
 Matches 140; Conservative 85; Mismatches 210; Indels 4; Gaps 2;

QY 1 MKPPPIORPTNFOAVAR---IGYGIATLTFVGLGMAAFPLDSAVANG-VYSAEVSQ 56
 Db 36 MRMNRIDORTVYSRSIRKHLIAGLAGALLAGVGMAATTNLGAVVASGHLVDSYTK 95

QY 57 DVOLLEGMLAKIIVREGEKAKGVLFELDPQTQANAAAGITRNQYVALKMEARLLAER 116
 Db 96 KVGHPKGVGEILLVNEGERKAGDVWRDLATQTRANLAVTRRLDELGARMARLEAER 155

QY 117 DQRSISPPADLTSGRADPMYARAIADBOAQFTERRQTIQGVDMNQRLOVSEIEGI 176
 Db 156 DDLAEIIPQALIRRDPPDVAASMRSETKLFEEFKSYREGRKQALERTIOFHEIEIGL 215

QY 177 DRGTGKLDQGFIEDELIDRLKLYDKGLVPRRLALEANAGSLSSIGRLTDRSKAV 236
 Db 216 KAQEVAYDNGLAIVLEAETTSOKSLREGVSVQRNLSQTQAAITFGGEREKIAYQOTA 275

QY 237 QGASDTOLKVRQIKOEFEQVSQSTITERVRLAEVTEKEVVASDAQRKIKIVSPVNGTAQ 296
 Db 276 GRITETKQIQLIDQELTEVGRLEIQAQMGFEVERKVAEDELKRIDIVAQSGMVH 335

QY 297 NLRFETGAVVRAAPLVLDIAPDEDAFYQAHFOPTDVNVHMGVTEVRLPAFHSAGNP 356
 Db 336 EMAVHTVGVVTPADPILVYPDGDLEALVQIVPKDIDQVQOKAMLRMTAFNQRTVP 395

QY 357 DPERHDPAVAVRISDPQKARLFGLVRYVYKOLPHLRNVYAGPAQVIVPGEETV 416
 Db 396 ELEGVASIAADITTDQRTGISTYLARISVPSEREKLNNAPLVGPMAEAFIQTSEKTA 455

QY 417 LQYLESPDLRLRTTMRRE 435
 Db 456 LSYIAKPLTDQISRAFREE 474

RESULT 3
 AE2972

secretion protein, HlyD family [imported] - Agrobacterium tumefaciens (strain C58, Dupont
 C:Species: Agrobacterium tumefaciens
 C:Date: 11-Jan-2002 #sequence.revision 11-Jan-2002 #text.change 11-Jan-2002

C:Accession: AE2972
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McCelli
 ; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

ster, E.W.
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:11743193
 A:Accession: AE2972

A:Status: preliminary
 A:Molecule type: DNA

A:Residues: 1-437 <KUR>
 A:Cross-references: GB:AE008689; PIDN:AAU44195.1; PID:917741773; GSPDB:GN00187

A:Experimental source: strain C58 (Dupont)
 C:Genetics:

A:Gene: prse
 A:Map position: linear chromosome

Query Match 27.0%; Score 584; DB 2; Length 437;
 Best Local Similarity 31.9%; Pred. No. 5.8e-29;
 Matches 139; Conservative 84; Mismatches 209; Indels 4; Gaps 2;

QY 4 PKIQPTNFOAVAR---IGYGIATLTFVGLGMAAFPLDSAVANG-VYSAEVSQ 59
 Db 2 PNIDORTVYSRSIRKHLIAGLAGALLAGVGMAATTNLGAVVASGHLVDSYTKKVQ 61

QY 60 HLEGMLAKIIVREGEKAKGVLFELDPQTQANAAAGITRNQYVALKMEARLLAERDOR 119
 Db 62 HPKGVGEILLVNEGERKAGDVWRDLATQTRANLAVTRRLDELGARMARLEAERDL 121

QY 120 PSISFPADLTSGRADPMYARAIADBOAQFTERRQTIQGVDMNQRLOVSEIEGIDRQ 179
 Db 122 AELIFPQALIRRDPPDVAASMRSETKLFEEFKSYREGRKQALERTIOFHEIEIGLAKQ 181

QY 180 TQGLKLDQGFIEDELIDRLKLYDKGLVPRRLALEANAGSLSSIGRLTDRSKAVGA 239
 Db 182 EVAVDNGLAIVLEAETTSOKSLREGVSVQRNLSQTQAAITFGGEREKIAYQOTAGRI 241

QY 240 SDTOLKVRQIKOEFEQVSQSTITERVRLAEVTEKEVVASDAQRKIKIVSPVNGTAQNL 299
 Db 242 TETKQIQLIDQELTEVGRLEIQAQMGFEVERKVAEDELKRIDIVAQSGMVHMA 301

QY 300 FTEGAVVRAAPLVLDIAPDEDAFYQAHFOPTDVNVHMGVTEVRLPAFHSAGNDPE 359
 Db 302 VHTVGVVTPADPILVYPDGDLEALVQIVPKDIDQVQOKAMLRMTAFNQRTVELE 361

QY 360 RHDPAVAVRISDPQKARLFGLVRYVYKOLPHLRNVYAGPAQVIVPGEETV 419
 Db 362 GHVSRIADITTDQRTGISTYLARISVPSEREKLNNAPLVGPMAEAFIQTSEKTA 421

QY 420 LFSPLRDLRTTMRRE 435
 Db 422 TAKPLTDQISRAFREE 437

RESULT 4
 C96003

protein secretion protein, HlyD family [imported] - Sinorhizobium meliloti (strain 10
 C:Species: Sinorhizobium meliloti
 C:Date: 24-Aug-2001 #sequence.revision 24-Aug-2001 #text.change 30-Sep-2001

C:Accession: C96003
 R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Her
 Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A:Title: The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-fixing e
 A:Reference number: A95842; MUID:21396508; PMID:11481431

A:Accession: C96003
 A:Status: preliminary
 A:Molecule type: DNA

A:Residues: 1-439 <KUR>
 A:Cross-references: GB:AL591985; PIDN:CAC49691.1; PID:915141178; GSPDB:GN00167

A:Experimental source: strain 1021, megaplasmid PSYMB
 R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubl
 pella, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.

L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau
 behault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,

A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.


```

A:Accession: H97683
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-437 <URL>
A:Cross-references: GB:AE007869; PIDN:AAK88425.1; PID:915157919; GSPB:GN001693
C:Genetics:
A:Gene: AGR_C_4904
A:Map position: circular chromosome

```

Query Match	25.1%;	Score 542.5;	DB 2;	Length 437;
Best Local Similarity	30.8%;	Pred. No. 2.2e-26;		
Matches 130;	Conservative 83;	Mismatches 204;	Indels 5;	Gaps 2

OY	15	A V A I G G I I L T F V G L G M A A P L D S A V I A N G V S A E - V S O V U H E G M K A K T I R E	73
Dd	20	A V V I L G I G L L - - - - V G M G N A F A F K L A G A V V A T G R V Y E G S K K I Q H L S G G I S V S I N V E	75
OY	74	G E K X K A G V L F E L D P T O A N N A A G I T R N O Y V A L K M E A R L L A E R D R P S I S P A D L T S O R A	133
Dd	76	G D R V A A Q O I L L R A T S A V V Q A N L S T I E N T L A O L Y S R R A L R A E I E E P S F T E D T L A T L S	135
OY	134	D P W A R A I A D E A Q O F T E R R O T I G O V D L M N O R L O Y S E I G I R O T O G L D O L G F I E D E	193
Dd	136	S K S A K T I D S Q N L F N S R R N L I G M K K O L A T R K I Q L A D E A N G L D V O E A T E M E I A Y K E D	195
OY	194	L I D R K I Y D K G L V E R P R L L A E A R A G S L S G S I G R L T A D S K A V O G A S D P O L K V R O I K O E F	253
Dd	196	V S K T D E L L K K G L V T Q R L N L K R O L S N L E G O G O Y I A B A Q T Y G K L S E L D Q L L Q D B E D R	255
OY	254	F E O Y S O S T T E R R V L A E V T E K E V Y A S D A Q R I K I Y S P N G N A Q N I R F T E G A V Y R A A P L	313
Dd	256	K S E Y T K D L T S I E A V T A E E E R L A A T R Q D L R D L R S P A G R I Y O L S V N I N G V I O P G E V L	315
OY	314	V D I A P E D E A F I O A H F O P T D V D N V M G M V T E V R L P A F H S A C N P D E R H D P A V A D R I E D P	373
Dd	316	M L V P D M D D L A I E A N I P R D I O I Y G O P V T F T A F N O S T T P D L S A V A V A Y A D L Q I D S	375
OY	374	Q K Q A R L E G I V A V D V K O L P P H L R G R V T A G M P A O V I V P T G E R T V L O Y L F S P L R D L T R T M R	433
Dd	376	R T G T S Y Y V L R I R P N K A G M G H L P G C K L Y P G M P A E V F I Q T S E S S V L Y E V K P R O D R L K T F V	435
OY	434	E E 435	
Dd	436	Q E 437	

RESULT 10

AI2908
H1YD family secretion protein [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont)
C:Species: *Agrobacterium tumefaciens*
C:Date: 11-Jan-2002 #sequence-revision 11-Jan-2002 #text-change 11-Jan-2002
C:Accession: AI2908
R:Ward, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
R:Maage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutlyavin, T.; Levy, R.; Li, M.; McDaniel
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AI2908
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-437 <KUR>
A:Cross-references: GB:AE008688; PIDN:AL43687.1; PID:917741213; GSPDB:GN00186
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: prsE
A:Map position: circular chromosome

Query Match	25.1%;	Score 542.5;	DB 2;	Length 437;
Best Local Similarity	30.8%;	Pred. NO. 2.2e-26;		
Matches 130;	Conservative 83;	Mismatches 204;	Indels 5;	Gaps 2

Oy	15	AVAAGVYIIALTFFGLTGMAAFAPLDASVANGVSAE-VSODVOHLEGGMKACILVRE	73
		: : : : : : : : : : : : : : : : : :	
Dd	20	AVVYVLGGLL-----VGNGMAAFAKLAGAAVAATGRVYVEGSKIKIHLSSGIYSBINYE	75
Oy	74	GKVYKAGVLEELDPQOANAAGITRNOYVALKAMEARLLAERDQPSISFPADLTSGRA	133
		: : : : : : : : : : : : : : : : : :	
Dd	76	GDRVAGAGCILLRLSATVVQAMNLSIENTLALQYSRRARLRAEIEEPPFTVEDTLTALS	135
Oy	134	DPMARATADDAQOTTEROTIQOGVDLMNMNRLOQYOSEIGIDRGQGLDLDOGFIDE	193
		: : : : : : : : : : : : : : : : : :	
Dd	136	SKSAKTFIDSQNLFNFSRNALLIGKKOLATRKLQLADEARGLDVQEATENELAIVED	195
Oy	194	LIDRLKLYDGSLVPRPRLALEARAGSLSGSIGRLTADRSKAVOGASTQLKVRQIOEF	253
		: : : : : : : : : : : : : : : : : :	
Dd	196	VSKTDDELKKGIVTLQIRNLMLKROLQSLNEGGOGVYIARAQTVCGLSTELPLQLLOJDBDR	255
Oy	254	FEOYSOSTETFRVRLAEVTEKEVVASDAQRKIKYSPNPGTAONURFTBEAVYBRAEP	313
		: : : : : : : : : : : : : : : : : :	
Dd	256	KSEVVKDKLTSTIEATVAAEEERIATAARDQLDIRDSIPAGRITYOLSVHNINGVLOPGSVL	315
Oy	314	VDIAPEADEAFYIOAHFOPTDVDNVMHGWATEVRLPAFHSACPNDPERHDPAVAVADRISP	373
		: : : : : : : : : : : : : : : : : :	
Dd	316	MUYPDKRDDLAIENAITTRDDIDQIYGOPVYVRFATFNQSTTPDLSAAYAAVADLODTS	375
Oy	374	OKOARLLGTVRVVDKQPLPHLRGRVTAGMPAQYIVPTGERTVIQYLFE SPLRDLTRMTMR	433
		: : : : : : : : : : : : : : : : : :	
Dd	376	RGTGSYYLVLRIRPNKAGMGHLPGGKLYPGMFAEVEFIQTSEKSVLSYFWPKPODRLRKTFV	435
Oy	434	EE 435	
Dd	436	OE 437	

RESULT 11

D971753
alkaline protease secretion protein AprE [Imported] - Rickettsia conorii (strain Mal11)
C:Species: Rickettsia conorii
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 30-Sep-2001
C:Accession: D97753
R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.;
Science 293, 2093-2098, 2001
A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A:Reference number: A97700; MUID:21442074; PMID:11557893
A:Accession: D97753
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-511 <KUR>
A:Cross-references: GB:AE006914; PIDN:AL02966.1; PID:g15619497; GSPDB:GN00173
C:Genetics:
A:Gene: aprE

Query Match	24.5%	Score 530.5	DB 2	Length 511
Best Local Similarity	30.0%	Pred. NO. 1.5e-25		
Matches 135; Conservative	88;	Mismatches 202;	Indels 25;	Gaps 7.

QY 5 KIORTDFMFAVAR--IYGIJALFEVILG--WAEFPLDSAVIANCVSAEVS-DOVQ 59

Db 68 KTDKORNNVQAARSPILFGIYVMFVLVLIGLMSALAPLDSGAVAVGIVMPSITKTKIQ 127

QY 60 HLEGMLAKILIVREGSEKYKAGVLEFELDPQANAAAGITRRQYALAKMEKRLAERDQR 119

Db 128 HHGEGIIIAIYVQGDKKVEDSKLEIETETIKSEHENILLOQYRNFLETENRLAERNDL 187

QY 120 PSISPADLTSGRADPWVARALADEQAQFTEBRQRTIQGVQVLMNARLOVYSEJEGIDRO 179

Db 188 EOIIFSDFLMODINLPEYAKITIIHQENLFRSRKEEYVNSKALHONIAOLEKTEGLEAK 247

QY 180 TQGLKDLQGFIEDELIDLRKLYDGLVPRPRLALLEARAGSLSGISGRITLADRSKAVOGA 239

Db 248 KVAASKTAEVYQDRKALARTLKEKGFVQKALLDQEAQVAAASKSDVATTEAKIINGIRHAI 307

QY 240 SDTOLTKAQIKOEFEFVOSISITETRRVRLAEVTEKEVVAQAKRIKIVSPVNGTAQNL 299

Db 308 TETQIKTIINQONKYTEERLTTELREAOVOTASIKERYNALTSJLNKVIIRAPDGVNSIK 367
 QY 300 FFTGAIVRAAEPLVDIAPDEDAFVIOAHFOPTVDVNMGMVTEVRLLPAFHSAGNP--- 356
 Db 368 YHTIGVISHQOPIMEISPTNDPLILEKAVSOKNIDSVHEGLVAKIRSAKSRTPPTT 427
 QY 357 -DPERHDPVAADRIISDF-QKQARLEGLIVND-----VKQLPRLGRVYAGMPA 405
 Db 428 GKVVSIISPDIVQDERQYGGQODNYVARVEIDMEFNKAVKKNLELH-----PGMOA 481
 QY 406 QVIVPTGERTLOYLFSPLRDTLRTMREE 435
 Db 482 EVQIVTGTTRLLRYLDDVTDPAFAKREK 511

RESULT 12

D1687
 alkaline proteinase secretion protein apre (apre) RP314 - Rickettsia prowazekii
 C:Species: Rickettsia prowazekii
 C:Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000
 C:Accession: D1687
 R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Slicheritz-Ponten, T.; Alsmark, U
 Nature 396, 133-140, 1998
 A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
 A:Reference number: A71630; MUID:99039499; PMID:9823893
 A:Accession: D1687
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-511 <AND>
 A:Cross-references: GB:AJ235271; GB:AJ235269; NID:93868717; PIDN:CAAL4774.1; PID:93868087
 A:Experimental source: strain Madrid E
 C:Genetics:
 A:Gene: apre; RP314

Query Match 24.0%; Score 519.5; DB 2; Length 511;
 Best Local Similarity 29.3%; Pred. No. 7.1e-25;
 Matches 134; Conservative 93; Mismatches 192; Indels 39; Gaps 10;

QY 6 IQRPD---NFQAVAR--IGYIATLVGLG--MAAPALDSAVANGVSAEVS-Q 56
 Db 65 ITRKTEDERRNVAAQARSPLFGIYVIFLVIGLMSALAPDSCGAVAGIVIPSTKK 124
 QY 57 DVQHLGGLAKIIVREBEKAKQAVLELDPQANAAAGITRNOYVALKMEARLLAER 116
 Db 125 TIQNEGGIINAIYKQDGKBEKDLIEETRIKSEHENILSQYRFLATENMLAER 184
 QY 117 DQRSISFPADLTQSRAD-PWVARAIADQEQFERRQTIQGVYDLMNAORLOYOSEIEG 175
 Db 185 DNLQKINF-SDFLMQNTINLPVAKIIHQENLFKSRKEVYSEKDAFQNTAOLEKIDG 243
 QY 176 IDROTQGLKQDLGFIEDLDLRLKYDKGLVPRRLALLEARAGSLSGISRLTADRSKA 235
 Db 244 LEAKRIAAKSTSEYQDRLKALFLKEGFIQKALLQEQAKVAASDAVTEAEIAGI 303
 QY 236 VQGSDDQKVRQKQKQEFEEVQSITETRVRLAEVTEKEVAVADDAQRKIVSPVNGTA 295
 Db 304 RHAITETQIKINHONKYTEERLTTELREAOVOTASIKERYNSITDNLNRIIRSPVDIV 363
 QY 296 QNLRFEGAVVRAAEPLVDIAPDEDAFVIOAHFOPTVDVNMGMVTEVRLLPAFHSAGN 355
 Db 364 NNLKVIHTIGVISHQOPIMEISPTNDPLILEKAVSOKNIDSVHEGLVAKIRSAKSRTP 423
 QY 356 P-----DPERHDPVAADRIISDPQKARLEGLIVND-----VKQLPRLGR 397
 Db 424 PTFGKVVSIISPD---IVDERQHLGQODNYVARVEIDMEFNKAVKKNLELH--- 476
 QY 398 RVTAGMPAQLVPTGERTLOYLFSPLRDTLRTMREE 435
 Db 477 ---PGMOAEVQIVTGTTRLLRYLDDVTDPAFAKREK 511

RESULT 13

S26697
 apre protein - Pseudomonas aeruginosa
 C:Species: Pseudomonas aeruginosa
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 31-Mar-2000
 C:Accession: S26697
 R:Duong, F.; Lazdunski, A.; Cami, B.; Murgier, M.
 Gene 121, 47-54, 1992
 A:Title: Sequence of a cluster of genes controlling synthesis and secretion of alkali
 A:Reference number: S26696; MUID:93051361; PMID:1427098
 A:Accession: S26697
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-432 <DUO>
 A:Cross-references: EMBL:X64558; NID:945279; PIDN:CA45856.1; PID:945281
 C:Superfamily: hemolysin secretion protein D; lipoyl/lipotin-binding homology
 C:Keywords: transmembrane protein
 F:48-91,284-313/Domains: lipoyl/lipotin-binding homology #status atypical <LPB>

Query Match 23.8%; Score 514.5; DB 2; Length 432;
 Best Local Similarity 31.6%; Pred. No. 1.2e-24;
 Matches 137; Conservative 88; Mismatches 182; Indels 27; Gaps 11;

QY 15 AVARIGVGIATLVTFVGLGMAAPALDS--AVANGVSAEVSQDVQHLGEGMLAKILVR 72
 Db 11 AYARGLMLVLFEGGALMLAPALDSGVAPATVITISGO-KRSVQHPDGAVKHLIVR 69
 QY 73 EGEKAKQAVLELDPQANAAAGITRNOYVALKMEARLLAERDQPSISFPADLTQSR 132
 Db 70 DQNVHVEGEPLIRREPTQARANVSLNRYANANRLQARLQAEVDGRTLEMPAGLAEQA 129
 QY 133 ADPVARAIADQEQFERRQTIQGVYDLMNAORLOYOSEIEGIDROTQG---LKQDLG 188
 Db 130 PLPTLGERLT-ELQROLHSROTALANELSLRANIEGLRAQLEGI-ROTEGNORLQRL- 186
 QY 189 FIEDELIDRLKLYDKGVRPRRLALE-----ARAGSLSGISGRRLADRSKAVQAGSD 241
 Db 187 -LNSQLSARLQAEVGYPRKQLEQERQLAEVNARRLSSESGRGQIR-----QSIAE 238
 QY 242 TQLKVRQIKQEFEEVQSITETRVRLAEVTEKEVAVSADQAKRIKIVSPVNGTAQNTLRF 301
 Db 239 AQMRIAQREEEYRREVNQQLAETQVNNARTLWEELSSAREYELRHAIRPAGVYAGLKFV 298
 QY 302 TEGAVVRAAEPLVDIAPDEDAFVIOAHFOPTVDVNMGMVTEVRLLPAFHSAGNPDEPRH 361
 Db 299 TDGGAIVGPELIMYIVPNSDSELEVGQALVNLVRIHSGLEVENLFTAFNSKPRPVATGE 358
 QY 362 DPVAVADRIISDPQKARLEGLIVRVYKQLPRLRG-RVTAGMPAQLVPTGERTLOYL 420
 Db 359 VTMVSADRLDEQKKQPYRVRQVDAAM-CKLKGIDIRGMVQVYVTRGERSLNTYL 417
 QY 421 FSPRLDTLRTMREE 434
 Db 418 FKPLFDRAHVALAE 431

RESULT 14

G83489
 alkaline proteinase secretion protein Apre PA1247 [imported] - Pseudomonas aeruginosa
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: G83489
 R:Stover, C.K.; Pham, X.Q.; Eyrin, A.L.; Micoquich, S.D.; Warrenner, P.; Hickey, M.J.;
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; L
 .; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
 A:Reference number: A82950; MUID:20437337; PMID:10984043
 A:Accession: G83489
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-432 <STO>
 A:Cross-references: GB:AE004554; GB:AE004091; NID:99947174; PIDN:ANG04636.1; GSPDB:GN
 A:Experimental source: strain PA01

C:Genetics:
A:Gene: apre; PA1247

Query Match 23.8%; Score 514.5; DB 2; Length 432;
Best Local Similarity 31.6%; Pred. No. 1.2e-24;
Matches 137; Conservative 88; Mismatches 182; Indels 27; Gaps 11;

15 AVARIIGYIALTFVGLGMAFAPLDS--AVIANGVSAEVSODVQHEGLMKLILVR 72
11 AVARIGMLLVLEFGGALLMAFAPLDGVANPATVITISG- RKSQVHGLGAVVHILVR 69
73 EGEKKAQGVLELDPQANAAAGITRNQYVALKMEARLARERDQPSISFPADLTSDR 132
70 DGOHVEASEPLTRMEPTQARAVNDSLNRNANARLQAEYDGRRTLEMPAGIAEQA 129
133 ADPMVARAIADEQAQFTERROT-IOGVDMNAQRLQOSELEGIDRQGV---LKDOLG 188
130 PLPTIGERL-ELQRLSLRSOTALNELSALNANIEGLRAQLEGL-RQTEGNORLQOURL- 186
189 FIEDELIDRLKLYDKGLVPRPRLALE-----ARAGSLSGSIGRLTADRKAQVAGASD 241
187 -LNSGLSGARDLAEGGYPRNQLLEORQLAEVNAKLSSSSRFGQIR-----QSTAE 238
242 TOLKVRQIKQEFFEQVQSITETRVRLAEVTEKEVVASDAQRIKIVSPNGCTAQNLEFF 301
239 AQMRIQGESEERYRKEVNGQLAETQVNAARTIMEELSSARELRAHEIRAPVSGYAGLKVF 298
302 TEGAVRRAERPLVDIAPDEAFVIOAHQPTVDVNMGMVTEVLLPFAHSGNDDPREH 361
299 TDGGVIGPEELMTVYPNSDSLVEBQLAVNLVDRHGLPEMLFTAPENOSKTRVYGE 358
362 DPAVAVADRISDPQKQARLEFLGIVRVVKQLPRLRG-RVTAQMPAQVIVPTGERTVLYL 420
359 VIMVADRLDDQNNQPYALRAQYDAAM-CKLKGLOIRPMANQVFPVTERGSLNTYL 417
421 FSPRLDFTTWMRE 434
418 FKPLFDRAHVALAE 431

RESULT 15

S12526
metallopeptidase export system membrane fusion protein prte [validated] - Erwinia chrys

C:Species: Erwinia chrysanthemi

C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 26-May-2000

C:Accession: S12526; T48667
R:Letoffe, S.; Delepelaitre, P.; Wandersman, C.

EMBO J. 9, 1375-1382, 1990

A:Title: Protease secretion by Erwinia chrysanthemi: the specific secretion functions at

A:Reference number: S12524; MUID:90228333; PMID:2184029

A:Accession: S12526

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-448 <LET>
A:Cross-references: GB:X53253; NID:941546; PIDN:CMA37343.1; PID:941549

R:Delepelaitre, P.; Wandersman, C.
submitted to the EMBL Data Library, April 1991

A:Description: C-terminal secretion signal in protease secreted by gram negative bacteria

A:Reference number: 224521

A:Accession: T48667
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-448
A:Cross-references: EMBL:M60395; PIDN:AAA63635.1

A:Experimental source: strain B374

C:Genetics:
A:Gene: prte

C:Function:

A:Description: involved in the export of metalloproteinases [validated, MUID:90228333]

C:Superfamily: hemolysin secretion protein D; lipoyl/biotin-binding homology

C:Keywords: membrane-associated protein; protein export

F:64-107,300-329/Domain: lipoyl/biotin-binding homology #status atypical.<LPB>

Query Match 23.4%; Score 506.5; DB 2; Length 448;
Best Local Similarity 28.8%; Pred. No. 3.8e-24;
Matches 128; Conservative 93; Mismatches 191; Indels 33; Gaps 7;

8 RPTDNFOAVARIIGYIALTFVGLGMAFAPLDSAVIANG-VSAEVSODVQHEGLGML 66
20 RASRDEERLRLRGWMLVLAFGGFLMLALPLDGVANQGVVVSNGNKVIQHOGGIV 79
67 AKIIVREGKKAQGVLELDPQANAAAGITRNQYVALKMEARLARERDQPSISFPA 126
80 DRIOVKDGRVAAQGVLLTLNNAVDAETTSSEGSGSQYDOLIAAREARLARERDQPSISFPA 139
127 DLTSGRADPMVARAIADEQAQFTERROTIOGVDMNAQRLQOSELEGIDRQGV---LKDOLG 186
140 RLTOARQREPMATIALDE-----DLRSQSGSLKLEIDVRSIDGLETS 185
187 LGFIE-----DELIDRLKLYDKGLVPRPRLALEARAGSLSGSIGRLTADR 232
186 LGALQKVMSSKQSEQATLSQGLRPLAADVYVRNKKLTERLFAQVSGSLAQTSGEV 245
233 SKAVQASDPTOLKVRQIKQEFFEQVQSITETRVRLAEVTEKEVVASDAQRIKIVSPN 292
246 GTRRDIOQOKLRIARQOQYDKVNSLSQVAKLNEVYSQREKADFNLANVQARAPYA 305
293 GTAONLRFETEGAVVRAERPLVDIAPDEAFVIOAHQPTVDVNMGMVTEVLLPFAHSGNDDPREH 352
306 GTVAVMKITFTEGGVIAIPQVMMDIVDEDPRLVDGRIPVEMVDKWSGLPVLQETAFESQ 365
353 AGNPDEPHERDPAVAVADRISDPQKQARLEFLGIVRVVKQLPRLRG-RVTAQMPAQVIVPTGERTVLYL 420
366 STTPRPVPGVITLLSDRLVD-EKDGTPYVGL-RIVQSEGRKSLGLETIKPMPVQGFVR 423
411 TGERTVLOYLSPRLDFTTWMRE 435
424 TGERSEFNYLFPKPLMDRMHIALTEE 448

Search completed: April 22, 2003, 12:32:47
Job time : 23 secs

THIS PAGE BLANK (USPTO)

PT polypeptide, useful for treating sewage, waste water and in the pulping process -

PS Claim 1; Page 35; 46pp; English.

CC The invention relates to a Caulobacter host cell for expression and secretion of a heterologous polypeptide. The Caulobacter host cell comprises at least one surface layer transport protein having an amino acid sequence homologous to the Caulobacter crescentus RsaD or RsaE proteins (AAB22786, AAB22787). The host cell further comprises a DNA construct comprising DNA encoding a polypeptide heterologous to a surface layer (S-layer) protein of the cell 5' from, and operably linked to, a DNA encoding a Caulobacter surface layer protein secretion signal, with the proviso that when the cell comprises transport proteins having the same sequence as both the RsaD and RsaE proteins, the secretion signal is not from the Caulobacter crescentus surface layer protein (RsaA). The invention also encompasses a method for identifying a Caulobacter suitable for use as a host cell for expression and secretion of a heterologous polypeptide comprising the detection of homologues of the rsad or rsae genes using a rsad/rsae hybridisation probe. The host cell is used for the expression and secretion of a heterologous polypeptide. The modified Caulobacter cells may be used to treat sewage and waste water. The cells may also be grown in wood pulp suspensions and then used in the wood pulping process. They can also be used in fish vaccines. Caulobacter species are able to form biofilms, and can attach themselves to surfaces without producing the extracellular enzymes or polysaccharide slimes that are characteristic of most other surface-attached bacteria. Caulobacter are therefore particularly suited for use in bioreactor systems. The present sequence represents the CC Caulobacter crescentus membrane fusion protein (MFP), RsaE. This protein is anchored in the inner membrane of Caulobacter crescentus and is thought to span the periplasm. Together with an outer membrane protein, CC it forms a channel that extends from the cytoplasm through the two membranes to the outside of the cell.

XX Sequence 435 AA;

SO Query Match 100.0%; Score 2165; DB 21; Length 435;
Best Local Similarity 100.0%; Pred. No. 3, 2e-178;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKPKRIQPTNFGQVAIAIGYITATFTVGLGMAAFPLDSAVIANGVSAEVSQYOH 60
DB 1 MKPKRIQPTNFGQVAIAIGYITATFTVGLGMAAFPLDSAVIANGVSAEVSQYOH 60
QY 61 LEGGMLAKIIVREGEKVACGVLELDPDQANAAAGITRNQYVALKAMEARLLAERDORP 120
DB 61 LEGGMLAKIIVREGEKVACGVLELDPDQANAAAGITRNQYVALKAMEARLLAERDORP 120
QY 121 SISPADLTSGRADPMVARAIADBOAQFTERRQTIQGVDDIMNAQRIQYQSEIEGIDROT 180
DB 121 SISPADLTSGRADPMVARAIADBOAQFTERRQTIQGVDDIMNAQRIQYQSEIEGIDROT 180
QY 181 OGILDOGLFEIDELIDLRKLYDKGVPRLRLALEARAGSISGSGRTTARSRKAVOGAS 240
DB 181 OGILDOGLFEIDELIDLRKLYDKGVPRLRLALEARAGSISGSGRTTARSRKAVOGAS 240
QY 241 DTOLKVRQIKOEFEEYOSISITETRVRLAEVTEKEVNASDAQRIKRIYSPNGTAQNIRF 300
DB 241 DTOLKVRQIKOEFEEYOSISITETRVRLAEVTEKEVNASDAQRIKRIYSPNGTAQNIRF 300
QY 301 FTGGAIVVAAEPLVDIAPEDEAFVIAHFQPTDVNVMHGMVTEVRLPAFHSAGNPDEP 360
DB 301 FTGGAIVVAAEPLVDIAPEDEAFVIAHFQPTDVNVMHGMVTEVRLPAFHSAGNPDEP 360
QY 361 HDPVAVADRISDPQKQALFLGIVRVYKQLPRLRGVATGMPAQVYVPPGERVLOYL 420
DB 361 HDPVAVADRISDPQKQALFLGIVRVYKQLPRLRGVATGMPAQVYVPPGERVLOYL 420
QY 421 FSPPLDRLTTMRRE 435
DB 421 FSPPLDRLTTMRRE 435

RESULT 2

AAW05298

ID AAW05298 standard; Protein; 443 AA.

XX AC AAW05298;

DT 30-DEC-1996 (first entry)

XX DE Esterase secretory protein (ORF2 product).

XX KW Esterase secretory gene; protein secretion.

XX OS Serratia marcescens strain S-r41 (FERM BP-487).

XX PN EP733707-A2.

PD 25-SEP-1996.

XX PF 20-MAR-1996; 96EP-0104429.

XX PR 23-MAR-1995; 95JP-0063772.

XX PA (TANA) TANABE SEIYAKU CO.

XX PI Akatsuka H, Kawai E, Shibata T;

XX DR WPI: 1996-444599/45.

XX DR N-PSDB: AAT39662.

XX PT New isolated esterase secretory gene from Serratia - used to increase the productivity and extracellular secretion of esterases

XX PS Claim 3; Page 21-22; 29pp; English.

XX 3 Different proteins (AAW05297-99) participate in the mechanism of secretion of esterase by Serratia marcescens S-r41. They are encoded by an esterase secretory gene (AAT39662) that includes 3 CC open reading frames. The proteins can be produced by transformed CC host cells, pref. S. marcescens or E. coli, carrying vector CC plasmids incorporating the esterase secretory gene. The esterase CC is secreted from the host cell, and is recovered from the culture CC medium and from within the cells. The esterase is useful in CC hydrolysis reactions.

XX Sequence 443 AA;

SO Query Match 24.1%; Score 521.5; DB 17; Length 443;
Best Local Similarity 29.3%; Pred. No. 1, 6e-36;
Matches 133; Conservative 88; Mismatches 194; Indels 39; Gaps 7;

QY 6 IORPTDNF-----QAVARIGYITATFTVGLGMAAFPLDSAVIANGVSAEVS-Q 56
DB 5 IGEQDSTYETIEPQDERFTFMGLVVGIGFLGMAAFPLDKGVASPSVTVSGNRK 64
QY 57 DVQHLLEGMLAKIIVREGEKVACGVLELDPDQANAAAGITRNQYVALKAMEARLLAER 116
DB 65 TVQAPASGIIKINIAVIRGDKVAGEVYVQLSOVQAQVDSLROQYVTTLATERLLAER 124
QY 117 DQRSISFPADLTSGRADPMVARAIADBOAQFTERRQTIQGVDDIMNAQRIQYQSEIEG 176
DB 125 DGLSIVTFSPILDVAVKDPVRAEITIALQTLFASRRQALQSEID-----GKQSMGDI 177
QY 177 DROTGKGLD-----OLGFEDELIDLRKLYDKGVPRLRLALEARAGSISGSGRLT 229
DB 178 RFOLKGLDQDSGNKNOIQLSLREOMNSKQALAGYILPRNRYLEVQAFEVNSIDETV 237
QY 230 ADRSKAVQASDTOLKVRQIKOEFEEYOSISITETRVRLAEVTEKEVNASDAQRIKRIY 289
DB 238 GRIGOLQKQLLESQQRIDQREFADYQREYRQLAQTDASDFRKKLQMADEDLGNTAITS 297
QY 290 PVNGTAONLREFTGCAVVAEPLVDIAPEDEAFVIAHFQPTDVNVMHGMVTEVRLPA 349
DB 290 PVNGTAONLREFTGCAVVAEPLVDIAPEDEAFVIAHFQPTDVNVMHGMVTEVRLPA 349

```

Db      298 PVDGIVGVLINFTQGGVVGAGDGLMDVVPQATILVVDRLKVDLPDKVYNGLPVDMFTA 357
Oy      350 FHSAGNDPDERHDVAVADRISD-----PQCARLFL---GIVRVQKLPPLHGRVTA 401
Db      358 FNQMTPTPIPGTVTLVLSADRLVDKANGCEPYIOMQTVTSPECKMLSGEDIKF----- 409
Oy      402 GMPAQVITPTGERTVLOYLFSPRLRDTLRTMTEE 435
Db      410 GMPVEVFVKTGSRSLSTLFEKPIIDRAHTSLTEE 443

RESULT 3
AA555920
ID      AA555920 standard; Protein: 444 AA.
AC
AA555920;
DT      15-FEB-2000 (first entry)
DE      Pseudomonas fluorescens ABC transporter cassette component B.
KM      ATP-binding cassette; transporter; operon; LipBCD; Serratia marcescens;
XX      microbe; protein secretion.
OS      Pseudomonas fluorescens.
JPI1276172-A.
XX      12-OCT-1999.
PD      27-MAR-1998; 98JP-0080597.
PR      27-MAR-1998; 98JP-0080597.
PA      (TANA ) TANABE SEIYAKU CO.
XX      MPI: 1999-626936/54.
DR      N-PSDB; AA222700.
XX
XX      An ABC transporter gene - used to increase the ability of a microbe to
PT      secrete proteins
XX
XX      Claim 2; Page 16-17; 28pp; Japanese.
XX
XX      This sequence represents an ATP-binding cassette (ABC) transporter
CC      component B from an Pseudomonas fluorescens strain 33 and encoded by
CC      an ABC transporter operon. The operon comprises 3 genes where the
CC      termination codon of the first gene overlaps with the initiation codon
CC      of the second gene. The sequences of the encoded proteins have 60, 44
CC      and 46 homology respectively to the LipBCD proteins from Serratia
CC      marcescens. The novel gene and protein can give or increase the
CC      ability of a microbe for secreting a protein.
XX
XX      Sequence 444 AA;
SO
Query Match 22.9%; Score 495; DB 20; Length 444;
Best Local Similarity 30.9%; Pred. No. 3.5e-34;
Matches 134; Conservative 86; Mismatches 201; Indels 12; Gaps 7;
Oy      9 PTDNQANVARIIGYGIATLFWGLG--WAAPAPLDSAVIANG-VYSAEVSQDVQHLBEG 64
Db      18 PEHGARFFARRGM--LLTVVGAGGFFLMASLAPLDQGIPOGTVVSGKRAVOTLSPG 74
Oy      65 MLAKLIVEEGEKVKAGVLFELDPQANNAAGITENOVYALKAMEARLLAERDORPSISF 124
Db      75 VYSRLVAGEGAVKOGQFLFRIDQTONADVHSLDAQYRAMASVARKMSERDNDOSTITF 134
Oy      125 PADLTSQRADPMVARALADEQAQFTERRQTIQGVYDLNNAORLYQYSEITEGIDROTGLK 184
Db      135 PAELT-SGNPDQALALVEGQRQLFESSREAFARQAGIRANTEGATQUNGMRARSDLT 193
Oy      185 DQGFIEBELDLKRLYKGLVPRRLLALAKRAGSLSGSIGRLTADSKAVQGSPTQL 244

```

Db	194	AAQASLRQDNLNQLADLNGITPRRNLMAYQKQLSVOODLQNGESGRVEGILESL	253
Oy	245	KVRIKIQEEFEQCSOSITETRVRLAEVTEKEVVASDAQRIKIVSPVNGTAONLRFETEG	304
Db	254	KLOOHSEEQKQEVRSQDLADQALRSLLTEQQLNSAGFDLQHSSEINAPADGIAVLSVHTFG	313
Oy	305	AVVRAAEPLVLDAPDEAEVIAHQPTPTVDVNHKMGVTEVRLPAFHSAGNPDPERHDV	364
Db	314	AVVRAGETLLEIVPDQTRLEVEGRULPVHLVDVGTLLPVDILFTAFNOSRTPRVGEVSL	373
Oy	365	AVADRIISPQKQARLFGLGVAVDVKQLP-PHLRGVY-TAGMAQVTVPPERTVLQYLS	422
Db	374	ISADQMDLDEKTMPPY--VLRITVSSSALEKHLGLVIRKFMRAEMFIRIGERSLWYLR	431
Oy	423	PLRDTLRRTMREE	435
Db	432	PLDRAGSALTEE	444
RESULT 4			
ID	AAV55923	standard: Protein; 443 AA.	
AC	AAV55923;		
DT	15-FEB-2000	(first entry)	
DE	Pseudomonas fluorescens ABC transporter cassette II component B.		
KM	ATP-binding cassette; transporter; operon; lipBD; Serratia marcescens;		
OS	Pseudomonas fluorescens.		
PN	JPI1276172-A.		
PD	12-OCT-1999.		
PF	27-MAR-1998; 98JP-0080597.		
PR	27-MAR-1998; 98JP-0080597.		
PA	(TANA) TANABE SEIYAKU CO.		
DR	WP1; 1999-626936/54.		
DR	N-PSDB; AA222701.		
PT	An ABC transporter gene - used to increase the ability of a microbe to		
PS	secrete proteins		
PS	Claim 8; Page 23-24; 28pp; Japanese.		
CC	This sequence represents an ATP-binding cassette (ABC) transporter		
CC	component B from an Pseudomonas fluorescens strain 33 and encoded by		
CC	the ABC transporter operon II. The operon comprises 3 genes where the		
CC	termination codon of the first gene overlaps with the initiation codon		
CC	of the second gene. The sequences of the encoded proteins have 63, 56		
CC	and 59% homology respectively to the lipBD proteins from Serratia		
CC	marcescens. The novel gene and protein can give or increase the		
CC	ability of a microbe for secreting a protein.		
XX			
XX	Sequence 443 AA;		
XX			
Query Match	22.1%;	Score 479.5; DB 20; Length 443;	
Best Local Similarity	29.0%;	Pred. No. 7,6e-33;	
Matches 130; Conservative	83;	Mismatches 196; Indels 39; Gaps	
Oy	9	PTDNFAV-----ARIGYGIILTFVGLGMAAFPLDSAVTANG-VVSAEVSODQHLE	62
Db	14	PVNNVALVLDKRYARLGMVLVYGAGFLGMALAPLDKGVAVSGKVMVSGHRKTYQHPS	73
Oy	63	GGMLAKILVRSGEKKAGQVLFELDPDTONAAAGITRNOYVALKMAEARLLAERQQRPSI	122

Db 74 GGIVRIVRGEVVSAGQVLIHLKETPLRCOMOSLRQSFIASQASEARLSNESEGLSAV 133
 QY 123 SEPADLTISQADPVMARAIADQAFERRQTIQGVDMNAQRLQYQSEIEGIDRQTG 182
 Db 134 VEGPELLN---EPEAAATLSIQRLQFSSRAQALA-----TEQGLFETTAGAQAQLRG 183
 QY 183 LKD-----QLGFIEDELIDRLKLYDKGLVPRPLALEARAGSLSSIGRLADRSKA 235
 Db 184 TREOSAKSVLQRTAMNEQLOGLRELARDGYIPRNRLLENERLYAQLDGAIADEFGRIQL 243
 QY 236 VQASADPQLQKQEFEEQVQSITETRYRLAEVTEKEVVASQAQRITVSVNTA 295
 Db 244 QROYELRLRLRQLGQEDQKRLQGLATETRRSDRLNRKLSAEPELLNSLARAPAAV 303
 QY 296 QNLRFTEGAVVRAAEPVLDIAPEDAEVIAHQPTVDVNHMGMTVEVRLPAFHSAGN 355
 Db 304 VGLDVTYEGVYIKPQQAQALMDIVPOGEPLVFARVYQVADKVRHPLPELLEFSARNOSTT 363
 QY 356 PDPERHDPVAVADRISDQKQARLF-----LGIVRVYKQLPPLHGRVYTAGMPAQV 407
 Db 364 PRVAGEVTLVSAADRVQDERTDEPYTLRAQVSAAGMRQLDGLQIRP-----GMPVEA 415
 QY 408 IVPTEGERTVLQYLFESPLRDTLRTTWRE 435
 Db 416 FVKTGERSMNLTKFPLLDRTTHMALVEE 443

RESULT 5

ABBS2753
 ID ABBS2753 standard; Protein: 415 AA.

XX AC ABBS2753;

DT 11-FEB-2002 (first entry)

DE Escherichia coli polypeptide SEQ ID NO 907.

XX KW Escherichia coli; B2/D+A-; antiinflammatory; antibacterial;
 KW immunosuppressive; extra-intestinal infection; phylogeny; meningitis;
 KW systemic infection; non-diarrhoeal infection; septicemia;
 XX KW pyelonephritis; antibiotic resistance.

XX OS Escherichia coli.

XX PN W0200166572-A2.

XX PD 13-SEP-2001.

XX PF 12-MAR-2001; 2001WO-EP03445.

XX PR 10-MAR-2000; 2000FR-0003145.

XX PR 02-FEB-2001; 2001FR-0001449.

XX PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

XX PI Blngen E, Bonacorsi S, Clermont O, Nassif X, Tinsley C;

XX DR WPI: 2001-550253/61.

XX PT A library of DNA fragments of Escherichia coli strains for the

XX PS phylogenetic determination of a given strain comprises polynucleotides of

XX PS nature B2/D+ A- -

XX PS Example 6; Fig 6; 646pp; English.

CC The invention relates to a library of DNA fragments of Escherichia coli
 CC strains comprising polynucleotides (ABA88577-ABA88729 and ABA89533)
 CC and encoded proteins (ABBS2459-ABBS2919 and ABBS2954-ABBS3094) of nature
 CC B2/D+A-. The polynucleotides have potential antiinflammatory,
 CC antibacterial and immunosuppressive activity as part of pharmaceutical
 CC compositions used to treat, palliate or prevent extra-intestinal E. coli
 CC infections. The polypeptides are useful for determining the phylogenic
 CC group of a given E. coli strain. These polypeptides can detect and treat

CC an undesired development of E. coli, particularly an extra-intestinal
 CC infection that include systemic and non-diarrhoeal infections such as
 CC septicemia, pyelonephritis and meningitis this is particularly
 CC advantageous as bacterial resistance is increasing with the more
 CC frequent use of broad spectrum antibiotics.

XX SQ Sequence 415 AA;

Query Match 15.5%; Score 335; DB 22; Length 415;
 Best Local Similarity 25.6%; Pred. No. 2e-20;
 Matches 110; Conservative 74; Mismatches 162; Indels 84; Gaps 11;

QY 23 IIALTFVGLCAAPAPIDSANVANG-VSAEVSODVOHLEGMLAKTLVREGKVAQ 81
 Db 47 VLVATFV---IWMANSPIDEVTRGQSTIPGSRQVITQTLDPGLIKLTVREGDIVKQ 103
 QY 82 VLFELDPQANAAGITNQNVALKAMEARLLAERDQPSISFPADLTISQADPVMARAI 141
 Db 104 VLLTIDTTRSSAMLRERARVNNLEAVARLARAE-YSESITFPDDV----- 149
 QY 142 ADEQAQFERRQTIQGVDMNAQRLQYQSEIEGIDRQTGKLDQGFIEDELIDRLKLY 201
 Db 150 ---PADLRESESTV-----YRLKTELQSIAGLKQSKALDKKEIAMTRPIV 193
 QY 202 DKGLVPRRLALEARAGSLSSIGRLTADRSKAVQASDTQLVKRQTKQEFEEQVQS 261
 Db 194 REGAMSEVELLRMQRSNEL-----QLQWDEKQNTYLTGAGAL 232
 QY 262 TETRYRLAEVTEKEVVASDAQKRITVSPVNGTQNLRFTEGAVVRAAEPVLDIAPDE 321
 Db 233 VKTEAELEQAQENAGNAGRPVPERSRIRAPLQYKNIIVNTLGGVVSAGQDIMEIPL 292
 QY 322 AFVIAHQPTVDVNHMGMTVEVRLPAFHS-----GNPDERHDPVAVADR 370
 Db 293 QLTIEAYINPRDAVYVFTGMPALVKLTAYDYAIVGGLDGVTVLVSPTLR-DQKRPGLK 351
 QY 371 SDPQKQARLFGIVRVYKQLPPLHGR-----VTAGMPAQVIVPTEGERTVLQY 424
 Db 352 LDPE-----AYYRVLTTSNNYLTDRNGKILPVPQMIASVDIKTGQKSVFYLKPI 405
 QY 425 RDTLRTTWRE 434
 Db 406 -----TRMKQ 410

RESULT 6
 AAY75573
 ID AAY75573 standard; Protein: 475 AA.

XX AC AAY75573;

XX DT 21-MAR-2000 (first entry)

XX DE Neisseria meningitidis ORF 764 protein sequence SEQ ID NO:2620.

XX KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
 XX KW antibacterial; gene therapy.

XX OS Neisseria meningitidis.

XX PN W09957280-A2.

XX PD 11-NOV-1999.

XX PF 30-APR-1999; 99WO-US09346.

XX PR 01-MAY-1998; 98US-0083758.

XX PR 31-JUL-1998; 98US-0094869.

XX PR 02-SEP-1998; 98US-0098994.

XX PR 02-SEP-1998; 98US-0099062.

XX PR 09-OCT-1998; 98US-0103749.

XX PR 09-OCT-1998; 98US-0103794.

09-OCT-1998; 98US-0103796.
PR 25-FEB-1999; 99US-0121528.
XX
XX (CHIR) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizsa M, Rappunli R, Ratti G, Scalato E, Scarselli M;
PI Tettelein H, Venter JC;
XX
XX WPI: 2000-062150/05.
DR N-PSDB: AAZ54335.
XX
XX
PT Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics -
XX
XX
PS Claim 2; Page 1243; 1453pp; English.
XX
XX AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
XX represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides
XX and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent
XX PCR primers used in the exemplification of the present invention. The
XX polypeptides, the polynucleotides, antibodies and compositions of
XX the invention can be used as vaccines, as diagnostic reagents, and as
XX immunogenic compositions. The polypeptides can be used in the
XX manufacture of medicaments for treating or preventing infection due to
XX *Neisserial* bacteria (e.g. meningitis and septicemia), to detect the
XX presence of *Neisseria* bacteria, or to raise antibodies. They may also
XX have use as antibacterial agents or antagonists, which may themselves
XX may also be used in gene therapy protocols.
XX
XX Sequence 4/75 AA;

Query Match	15.28;	Score 329;	DB 21;	Length 475;
Best Local Similarity	26.28;	Pred. No. 8e-20;		
Matches 128;	Conservative	188;	Indels	94;
			Gaps	18;

```

OY      1 MKPKRIORPDNF-----QAVRIRIGYIIALFVGLGLMAAFAPILDSAV 44
Db      27 LKPKKRAEDQAFPLPHLELTDPVPSAKRMAARF---IMARLLALL-WSMFKIDIVA 82
OY      45 IANG-VVSAEVSODVOHLEGGMLAKILVREGKVKAGOVLELDPF-----QANA 93
Db      83 AASKTVSGSKSKTIOPLTAVVYKAVHYRBDGHVQKGETLAEAVGTSDVVOSEQALQ 142
OY      94 AAGITRNOYVA-LKAMARLLAREDRPSTSPFADLTSGRADPMVARIALADQAOETER 152
Db      143 AAOLSKRTEYEAVALAESR-----TVPHIDMAQKRSIGLSADAVQSAQ 185
OY      153 QTIQGVLDLNAORLOAQOSEIEGIDRQTOGLKQO-----LGFIEDE-LIDLRKLYDGKL 205
Db      186 VLAHQYQAWAADAQAOQSALRGHQAELDSAKNQEQKIVSVCALIEQKADRYRLRLADNF 245
OY      206 VPRRLALAEKAGS-----LSGSIGRLTADRSKAAVGCASDLOKRAQIKOEFFE---QVS 258
Db      246 ISEHAFLEQOQSVSMNMNDLESTRGQROQIAIAQAEONRVLNTONLKRDLTDLRLQAN 305
OY      259 QSIIETFEVRLAEVTEKEVEVNASDAQKRIKIVSPVNGTAQOILREFTEGAAVVAERAPLVDIAP 318
Db      306 EQIDQYHGO-----TDK---AKORQOIMLTIOSPADRGIVQGEIATYTVGQVQAQAKMNVLAR 358
OY      319 EDEAFVYIOAHFOPLTDVNDVNHMGVTEYRLPAF-----HSAGNPDEERHDPVAVADRISD 372
Db      359 DDDKMDVDEVLVNLKNDIGFVQGDODAVVAKIESPFYTRYGVLTGKVSXSHDAVS----- 411
OY      373 POKARLEFLGIVRDVYKOLRPPH---LRGR---VTAGMPQOVIVIPNGEFTVLYOLFSPLRD 426
Db      412 -HBO-----LGLVYTAVVSLDKHTLNTIDGKRAVNLTAQMNVTAEIKGCKRRVLDVLLSPLOT 466
OY      427 TLRITMRE 434
Db      467 KLDSEPRE 474

```

Accession	Protein	Length	Score	E-value	Ident	Positives	Gaps	Ident	Positives	Gaps
AAW22161	standard; Protein; 477 AA.	477	100.00	0.00	477	477	0	477	477	0
AAW22161		477	100.00	0.00	477	477	0	477	477	0
AAW22161		477	100.00	0.00	477	477	0	477	477	0
16-FEB-1998	(first entry)	477	100.00	0.00	477	477	0	477	477	0
ApxiIID protein.		477	100.00	0.00	477	477	0	477	477	0
RTX toxin; apxICA gene; apxiIBD gene; apxiIAB/C gene; apxiIIABCD gene; repeat in toxins toxin; cell-associated RTX toxin; vaccine production; therapy; A. pleuropneumoniae infection; swine pleuropneumonia.		477	100.00	0.00	477	477	0	477	477	0
Actinobacillus pleuropneumoniae.		477	100.00	0.00	477	477	0	477	477	0
CA2170839-A.		477	100.00	0.00	477	477	0	477	477	0
02-SEP-1996.		477	100.00	0.00	477	477	0	477	477	0
01-MAR-1996;	96CA-2170839.	477	100.00	0.00	477	477	0	477	477	0
01-MAR-1995;	95US-0396244.	477	100.00	0.00	477	477	0	477	477	0
(UYGU-) UNIV GUELPH.		477	100.00	0.00	477	477	0	477	477	0
MacInnes J, Mallard B, Ricciolatti P, Rosendal S;		477	100.00	0.00	477	477	0	477	477	0
WPI; 1997-245536/23.		477	100.00	0.00	477	477	0	477	477	0
N-PSDB; AAT73220.		477	100.00	0.00	477	477	0	477	477	0

Disclosure; Pages 113-114; 151pp; English.

CC AAM221.15-W22161 represent *A. pleuropneumoniae* RTX (repeat in toxins).
CC CC toxins. This sequence are encoded by the aprXaA, aprXbD, aprXaIb7C, and
CC aprXIIIABCD genes (see AAT3217-173220), and can be expressed by
CC microorganisms used in the preparations of the invention. The
CC preparations are bacterial preparations comprising one or more isolated
CC and purified strains of a microorganism that produces one or more RTX
CC toxins, where the strains have at least one cell-associated RTX toxin.
CC The preparations are used for production of vaccines for the prophylaxis
CC and treatment of infectious diseases caused by microorganisms that
CC produce RTX toxins, where the strains have been attenuated or
CC inactivated. The vaccines are preferably against *Actinobacillus*
CC *pleuropneumoniae* infection (swine pleuropneumonia). It has been found
CC that *A. pleuropneumonia* produces significant quantities of
CC cell-associated RTX toxins when cultured under certain conditions, and
CC that the whole-cell protein composition of the cultures corresponds to
CC the whole-cell protein profiles obtained from cells recovered at
CC necropsy from the pleural fluid of infected swine. Vaccination with a
CC bacterin prepared from heat-inactivated cultures having significant
CC quantities of cell-associated RTX toxins give significant protection of
CC swine against challenge with homologous strains.

SQ Sequence 477 AA;

Query Match	14.5%	Score 313.5	DB 18	Length 477
Best Local Similarly	25.3%	Pred. No. 1.8e-18		
Matches 112; Conservative	88;	Mismatches 190;	Indels 53;	Gaps 13;

Qy	19	IGGIATFEGLLGMAFAPLDSAVIANG--VSAEVSODQHOHEGMLAKIIVREBEKV	77
Db	60	IATLMFLFAIY-ISTISVEIVASATGCTGSGSKETKPLFNALVDIYKQDFV	118
Qy	78	KAGOVLFEL-----DPTQANAAGITRNOYVALKAMEARLLAERDQPSISF-PADLT	129
Db	119	EKGQILLNLTAAGCDADQKRYVSLGIERLDGGRYKISLTSI--EHRRLPLFQNDQDFD	176

QY 130 S-QRADPVARAIAD-----QAQFERRQTTGGVDYDLMNAQRLQYQSEIEGIDR 178
 DB 177 SVOEDKTKGARHLITEQFETWOKRYOKELAVORKOAEKQVLAIRKYESASRIE---- 232
 QY 179 QTQGLKDQGLFTEDELIDRLKLYDKGLVPRRLALLEAR-----ASLSGSIGRLTADRSK 234
 DB 233 -----KELSLDKRLKYDVKSISKHELLAOKENRYEASNELSVYQSHKEVESD 280
 QY 235 AVQASDPTQLKVRQIKQEFEEQVSOSITETRYRLAEVTEKEVYASDAQRIK--IVSPVN 292
 DB 281 LKQAEDELKLYQLFKSPDLIEKIQNIREKQLTLELEKNE-----QROLASITIRAPVS 334
 QY 293 GTAQNLRFETGAVYRAAEPLVDIAPDEBAFYQAHFOPTVDVNHMGVTEVRLPAFHS 352
 DB 335 GTVQDLKTHRTKGVTYTAETLAVIAPEDDYLEVSALIONKDVGEVIEQEAIVKETEFPY 394
 QY 353 AGNPDPERHDPVAVADRISDPQKARFLGIVRVKOLPRHLRG-RVTAGMPAQVIPT 411
 DB 395 TRYGLYGVKVTITLDAIEHPQ-LGLVFNSTIEINKKTLTGDKKEIQSGMSVIAETKT 453
 QY 412 GERTVLYQLFSPRLDTLTMTRE 434
 DB 454 GERSVIFSLSPLESITESLRE 476

RESULT 8
 AAY51414
 ID AAY51414 standard; protein; 477 AA.

XX AAY51414;

DT 05-MAY-2000 (first entry)

DE A. pleuropneumoniae apxIIID protein.

XX RTX toxin; vaccine; Repeats in Toxins; immunostimulatory; antimicrobial;
 KW antiinflammatory; antirheumatic; antidiarrhoeal; treatment; pneumonia;
 KW pleuropneumonia; septicemia; nephritis; arthritis; endocarditis;
 KW shipping fever; abortion; whooping cough; sleepy foal disease; joint ill;
 KW urinary infection; peritonitis; meningitis; gastroenteritis;
 KW passive immunization; apxIIID.

XX Actinobacillus pleuropneumoniae.

XX USG019984-A.

XX 01-FEB-2000.

XX 23-DEC-1996; 96US-0772270.

XX 01-MAR-1995; 95US-0396244.

XX (UUGD-) UNIV GUELPH.

XX Mallard B, Rosendal S, MacInnes J, Ricciattelli P;

XX WPI; 2000-146864/13.

XX N-PSDB; AAZ88587.

PT Bacterial preparation comprising microorganisms which produce a member
 PT of the Repeats in Toxins (RTX) family, useful for treating swine
 PT pleuropneumonia, arthritis in swine, shipping fever and abortion in
 PT cattle, and sleepy foal disease -

XX Disclosure; Column 89-98; 96pp; English.

XX This invention describes a novel bacterial preparation (I) which
 CC comprises one or more isolated and purified strain(s) of a microorganism,
 CC cultured in tryptone yeast extract (TYE) broth, which produces one or
 CC more RTX toxins (belonging to the family of toxins referred to as Repeats
 CC in Toxins), where the strain(s) have at least one RTX toxin which is
 CC cell-associated. The products of the invention have immunostimulatory,

CC antimicrobial, antiinflammatory, antirheumatic and antidiarrhoeal activity.
 CC The bacterial preparation may be used as vaccines for the prophylaxis and
 CC treatment of infectious diseases caused by strains of microorganisms
 CC which produce one or more RTX toxins. The infectious diseases are swine
 CC pleuropneumonia, pneumonia, septicemia, nephritis and arthritis in swine;
 CC septicemia, nephritis, endocarditis and arthritis in piglets; shipping
 CC fever and abortion in cattle; whooping cough, sleepy foal disease or
 CC joint ill (purulent nephritis, arthritis) in foals; septicemia,
 CC polyarthritis and abortion in horses; and urinary infections,
 CC peritonitis, meningitis, and gastroenteritis. The bacterial preparations
 CC may also be used to prepare antibodies which may be used as a means of
 CC passive immunization. This sequence represents the Actinobacillus
 CC pleuropneumoniae apxIIID protein described in the method of the
 CC invention.

CC Sequence 477 AA;

Query Match 14.5%; Score 313.5; DB 21; Length 477;
 Best Local Similarity 25.3%; Pred. No. 1.8e-18;
 Matches 112; Conservative 88; Mismatches 190; Indels 53; Gaps 13;

QY 19 IGVGIIALTFVGLGMAFAPLDSAVIANG-VSAEVSQDVQNHLEGMLAKIIVEGKV 77

DB 60 IAYLMLFLFLAIY-ISIISKEIYASATGKIVGSGHKEIKPIENALVKDIPVQDGFV 118

QY 78 KAGQVLFEL-----DPTQANAAAGITRNOYVALKMEARILAEKQPSISF-PADLT 129

DB 119 EKGQILLNLTLGCDADKQKTVSLGLERLDGYRYKSLYSI--EHNRLPLDFFNOADFD 176

QY 130 S-QRADPVARAIAD-----QAQFERRQTTGGVDYDLMNAQRLQYQSEIEGIDR 178

DB 177 SVOEDKTKGARHLITEQFETWOKRYOKELAVORKOAEKQVLAIRKYESASRIE---- 232

QY 179 QTQGLKDQGLFTEDELIDRLKLYDKGLVPRRLALLEAR-----ASLSGSIGRLTADRSK 234

DB 233 -----KELSLDKRLKYDVKSISKHELLAOKENRYEASNELSVYQSHKEVESD 280

QY 235 AVQASDPTQLKVRQIKQEFEEQVSOSITETRYRLAEVTEKEVYASDAQRIK--IVSPVN 292

DB 281 LKQAEDELKLYQLFKSPDLIEKIQNIREKQLTLELEKNE-----QROLASITIRAPVS 334

QY 293 GTAQNLRFETGAVYRAAEPLVDIAPDEBAFYQAHFOPTVDVNHMGVTEVRLPAFHS 352

DB 335 GTVQDLKTHRTKGVTYTAETLAVIAPEDDYLEVSALIONKDVGEVIEQEAIVKETEFPY 394

QY 353 AGNPDPERHDPVAVADRISDPQKARFLGIVRVKOLPRHLRG-RVTAGMPAQVIPT 411

DB 395 TRYGLYGVKVTITLDAIEHPQ-LGLVFNSTIEINKKTLTGDKKEIQSGMSVIAETKT 453

QY 412 GERTVLYQLFSPRLDTLTMTRE 434

DB 454 GERSVIFSLSPLESITESLRE 476

RESULT 9

AAW22154
 ID AAW22154 standard; Protein; 478 AA.

AC AAW22154;

DT 16-FEB-1998 (first entry)

DE ApxID protein.

XX RTX toxin; apxICA gene; apxIIB gene; apxIIAB C gene; apxIIABCD gene;
 KW repeat in toxins toxin; cell-associated RTX toxin; vaccine production;
 KW therapy; A. pleuropneumoniae infection; swine pleuropneumonia.

XX Actinobacillus pleuropneumoniae.

XX CA2170839-A.

XX 02-SEP-1996.

[illegible]

Y 399 --VTGMPAQVYPTGERTVLQYFSEPLRDLRTTME 434
: :|||: ||||: | ||| : : :|||
Db 439 EIEI GSGMSVTAIEIKTGERSVISTLSPLEESVESLRE 477

RESULT 10
ID AAY51408 standard; protein: 478 AA.
AAY51408
AAVS1408;
05-MAY-2000 (first entry)

A. pleuropneumoniae clydD protein.

RTX toxin; vaccine; Repeats in Toxins; immunostimulatory; antimicrobial;
antiinflammatory; antiarthritic; antibovine; treatment; pneumonia;
pleuropneumonia; septicemia; nephritis; arthritis; endocarditis;
shipping fever; abortion; whooping cough; sleepy foal disease; joint ill;
urinary infection; peritonitis; meningitis; gastroenteritis;
passive immunization; clydD.

Actinobacillus pleuropneumoniae.

US6019984-A.
01-FEB-2000.
23-DEC-1996; 96US-0772270.
01-MAR-1995; 95US-0396244.
(UYGU-) UNIV GUELPH.

Mallard B, Rosendal S, MacInnes J, Ricciatti P:
WPI: 2000-146864/13.
N-PSDB: AAZ88585.

Bacterial preparation comprising microorganisms which produce a member
of the Repeats in Toxins (RTX) family, useful for treating swine
pleuropneumonia, arthritis in swine, shipping fever and abortion in
cattle, and sleepy foal disease -

Disclosure: Column 59-62; 96pp: English.

This invention describes a novel bacterial preparation (I) which
comprises one or more isolated and purified strain(s) of a microorganism,
cultured in tryptone yeast extract (TYE) broth, which produces one or
more RTX toxins (belonging to the family of toxins referred to as Repeats
in Toxins), where the strain(s) have at least one RTX toxin which is
cell-associated. The products of the invention have immunostimulatory,
antimicrobial, antiinflammatory, antiarthritic and antibovine activity.
The bacterial preparation may be used as vaccines for the prophylaxis and
treatment of infectious diseases caused by strains of microorganisms
which produce one or more RTX toxins. The infectious diseases are swine
pleuropneumonia, pneumonia, septicemia, nephritis and arthritis in swine;
septicemia, nephritis, endocarditis and arthritis in piglets; shipping
fever and abortion in cattle; whooping cough, sleepy foal disease or
joint ill (purulent nephritis, arthritis) in foals; septicemia,
polyarthritis and abortion in horses; and urinary infections,
peritonitis, meningitis, and gastroenteritis. The bacterial preparations
may also be used to prepare antibodies which may be used as a means of
passive immunization. This sequence represents the Actinobacillus
pleuropneumoniae clydD protein described in the method of the invention.

Sequence 478 AA:

Query Match 12.1%; Score 262; DB 21; Length 478;
Best Local Similarity 24.2%; Pred. No. 4.9e-14;
Matches 111; Conservative 86; Mismatches 178; Indels 84; Gaps 18;

19 IGYGIALTENGGLCAWFAPADLSAVIANGVVS-AEVSDVDHLEGGMIAKLVREGKEV 77

```

Db 60 IAYLMFLFALV-ISTVSHVEIYATATGKIASDSRKEKIPINALVKELFVODGQV 118
QY 78 KAGOVLELDPQANAAAGITRNOYVALKMEAR--LLAE--RDQPSISEPAD--LTS 130
Db 119 EKDOLLHLTALGADADQOKTSSLSLTKLEHYREYILLEVAADRLLIETTKDFKHA 178
QY 131 QRAPPMVARAIADBOAQTER-----KQTQGVADLMAQRLOYSIESIDRQT 180
Db 179 TEEDTRIRYLTTEQEFEMOKOKYOKELALORREAEKQTVLANIRK-----YEGISR-- 230
QY 181 QGLKQDLFEIDE-LIDRKLYDKGLVPRPRLALEAR---AGSLSGSIGRLTRADRSKA 235
Db 231 -----VENERLKDKLIFNSKSTSKHDVLTOENHIEAVNLAAYKSKRLNVESDL 281
QY 236 VQASDQOLKVRQIKQEFQVOSITETRVRLAEVTEKEVVAADQRIKIV--SPVNG 293
Db 282 RQAKEEHLITQLFRADILEKIKQNV-EAKQOLSLERK-----NEQKQIASVIRAPVSG 335
QY 294 TAONLRFTEGAVYRAAEPLVDIAPEDAEFYIOAHFQPTDVNVMGMVTEVRLPAFISA 353
Db 336 TVQOLKHTHTVGAVVTTAETLWVIAEDVDLEVTALIQNKDGIEVGODAVIKVETF-- 392
QY 354 GNPDERHDPVA-----VADRIQDQOKARLFLGIVRVYDVKQLPRLRGR--- 398
Db 393 -----PYTRGYLGMKKYKNTLEAIEHPQ--LGLVFNSIISIDRKT-----LSGKDGK 438
QY 399 ---VTAGMPAOVIVPTGERIVLQYLFSPLRDLTFTTMR 434
Db 439 EIELSGCMVTAETKGTGERSVISYLLSPLESVSESLRE 477

```

RESULT 11

AAV75574

ID AAV75574 standard; Protein; 435 AA.

AAV75574;

DT 21-MAR-2000 (first entry)

DE Neisseria meningitidis ORF 764 protein sequence SEQ ID NO:2622.

XX Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
 KW antibacterial; gene therapy.

XX Neisseria meningitidis.

PN WO957280-A2.

XX 11-NOV-1999.

PF 30-APR-1999; 99WO-US09346.

PR 01-MAY-1998; 98US-0083758.

PR 31-JUL-1998; 98US-0094869.

PR 02-SEP-1998; 98US-0098994.

PR 02-SEP-1998; 98US-0099062.

PR 09-OCT-1998; 98US-0103749.

PR 09-OCT-1998; 98US-0103794.

PR 09-OCT-1998; 98US-0103796.

PR 25-FEB-1999; 99US-0121528.

PA (CHIR) CHIRON CORP.

PI (GENO-) INST GENOMIC RES.

XX Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M,
 PI Petersen J, Piza M, Rappoli R, Ratti G, Scalato E, Scarselli M,
 PI Tettelin H, Venter JC;
 XX MPI; 2000-062150/05.
 DR N-PSDB; AA254336.
 XX

PT Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics
 PS Claim 2; Page 1244; 1453pp; English.
 CC AA53015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941
 CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
 CC and polypeptides. AA254537 to AA254576 and AA254616 to AA254673 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of
 CC the invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
 CC presence of Neisseria bacteria, or to raise antibodies. They may also
 CC be used to screen for agonists or antagonists, which may themselves
 CC have use as antibacterial agents. The polynucleotides of the invention
 CC may also be used in gene therapy protocols.

Sequence 435 AA;

Query Match 12.1%; Score 261.5; DB 21; Length 435;

Best Local Similarity 25.2%; Pred. No. 4.7e-14; Matches 102; Conservative 70; Mismatches 163; Indels 69; Gaps 14;

```

QY 23 IIALFVGLGMAFAPLDSAVIANG-VSAEVSODVOHLEGMLAKLVREGEKVRAGQ 81
Db 62 IMAFALLALIL-MSWFGKIDIVAAASGKTVSGRSKTIQPLETVVAVKAVHVRDGOHVXGE 120
QY 82 VLFEIDPT-----QANAAAGITRNOYVA-LKAMEARLLAERDQPSISEPADLTS 130
Db 121 TLAELEAVGTSDVYVOSKOLQAOALSRLRYEVALALESR-----TV 163
QY 131 QRAPPMVARAIADBOAQTERROTQGVADLMAQRLOYSIESIDRQTLGKQD--- 186
Db 164 PHDMAQARSLGSDADQVQSAQVLAHQYQAWAQQDQALGQALGHQAEQSAQAQEKXL 223
QY 187 --LGFIDE-LIDRKLYDKGLVPRPRLALEARAGS-----LSSISGRLTRADRSKANVQGA 239
Db 224 VSGAIEQOKTADYRRLRADNFIEBHAELEQOOSKSVSWNDLESTRGOMROIQAALQAE 283
QY 240 SDTOLKVRQIKQEFPE--QVSQITETRVRLAEVTEKEVVAADQRIKIVSPVNGTAQ 296
Db 284 QNRVLNQNLRDLIDLALQANEOIDQYRQ---TDK---AKORQQLMTIQSPADGTQV 336
QY 297 NLRFTEGAVYRAAEPLVDIAPEDAEFYIOAHFQPTDVNVMGMVTEVRLPAF----- 350
Db 337 ELATYTVGGVQAQKMMVVAAPDDDKMDEVVLVINKDIGFEVQGDADVVKIESPPYRNG 396
QY 351 HSAGNPDERHDPVAVADRISDPQOKARLFLGIVRVYDVKQLPPI 394
Db 397 YLTGKVKSVSHDAVS-----HEQ-----LGLVYTVAVSLDKH 428

```

RESULT 12

AAV24069

ID AAV24069 standard; peptide; 425 AA.

AAV24069;

DT 09-SEP-1999 (first entry)

DE Salmonella typhimurium peptide sequence SEQ ID NO:38.

XX Salmonella; Salmonellosis; detection; diagnosis; infection; food.
 KW Salmonella typhimurium.

OS Salmonella typhimurium.

PN US5925522-A.

XX 20-JUL-1999.

PF 09-MAY-1997; 97US-0853659.

Db 462 GGFAGRVEIATAPEPPAD---AALASEPLRYRRLA---GDAALRAGH-AAVLRPG 513
 QY 403 MPAGVPTGERTVQLQYLFSPRLDRLRTTMR 433
 Db 514 MRVGTALAEWRRESQMAFEPL-SSLHGTLR 543

RESULT 14

AA98387
 ID AAG98387 standard; Protein; 355 AA.

AC AAG98387;

DT 21-SEP-2001 (first entry)

DE Escherichia coli protein sequence SEQ ID NO:435.

KW Escherichia coli; identification; proliferation; microorganism;

KM antimicrobial; antibacterial; antibiotic; gene therapy; diagnosis;

KW bacterial growth inhibition.

OS Escherichia coli.

PN W0200148209-A2.

PD 05-JUL-2001.

PE 19-DEC-2000; 2000MO-US34419.

PR 23-DEC-1999; 9905-0173005.

PA (ELIT-) ELITRA PHARM INC.

PI Forsyth RA, Ohlsen KL, Zyskind JW;

DR WPI; 2001-457376/49.

DR N-PSDB; AAH81443.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Db 14 GLIYAV--AAIYAMLLRPAGVEGEFAVSNRIEA-TEVDIASKIAGRITLIVKEGFVR 70
 QY AGOVLFEIDPTQANAAAGITRNOYVALKAMBARLAEPRDQPSISFPADLTSGRADPMVA 138
 Db 71 EGEVIAKMD-----TR-----VLQEQRLAEIAQKENGQ----- 99

QY 139 RAIDDEQAQFERRQTIQGOVLDMAAQLQYQSEIEGI-DRQGLK-DQLGFIEDLID 196

Db 100 -AVAAQALLERQSEFRQAQSLVN---QRAEIDSVAKRHRTSRSLAQGAISAQOLD 154

QY 197 LRKLYDKLVPRRLALAEARAGSLSGISGLT/DRSKAVAGASDTQLKVRQIQEFEEQ 256

Db 155 -----DRAAASARAALESAKQVSAKAAIEARTNIID----- 190

QY 257 VSOSITETRVRLAEVTEKEVAVASDAOKRIKIVSPVNGAONLRFPTGCAVYRAAEPVLDI 316

Db 191 -----AOTRVAAQATERIADIDDSGLK--APRDGRVQ-YRAAEGEVIAAGRVLYNM 242

QY 317 APEDEAFYIQAHFOPTD-VDNVHMG-----MYTEVRLPA---FHSAGNPPPERHDP 363

Db 243 V--DLSDVYMTFFLPTEQAGTLKGEARLILDAAPDLRIPATISFVASVQAQFPK---T 297

QY 364 VAVADRISDPQKARLFLGIYRVGVKQLPRLRGRTVAGMAQYIVPGE 413

Db 298 VETSD-----ERLKMFRVKARIPPELLQGLE-YVXTGLPGVAMRVNE 341

RESULT 15

ID AA091064 standard; Protein; 412 AA.

AC AA091064;

DT 05-JUN-2002 (first entry)

DE Neisseria cell surface polypeptide #2.

XX Cell surface protein; antibacterial; antimicrobial.

XX Neisseria meningitidis.

XX W0200216612-A2.

XX 28-FEB-2002.

PE 21-AUG-2001; 2001MO-GB03759.

PR 24-AUG-2000; 2000GB-0020952.

PA (MICR-) MICROSCIENCE LTD.

PI Lane JD, Hughes WG, Santangelo JD;

DR WPI; 2002-280941/32.

DR N-PSDB; ABK54080.

XX Novel peptide encoded by Neisseria meningitidis, useful for manufacture

XX of medicament for treatment or prevention of condition associated with

XX infection by Neisseria or Gram-negative bacteria -

XX Claim 5; Page 19-20; 79pp; English.

XX The invention relates to polypeptides located on the cell surface of

XX Neisseria meningitidis, and the polynucleotides encoding them. The

XX sequences of the invention are useful for therapeutic or diagnostic use,

XX in the manufacture of a medicament for use in treatment or prevention of

XX bacteria. The sequences are also useful for screening potential

XX antimicrobial drugs or for detection of virulence. Sequences

XX AA091063-AA091079 represent Neisseria meningitidis polypeptides of the

XX invention.

XX Sequence 412 AA;

THIS PAGE BLANK (USPTO)


```

Db 439 HI--PLSSGMAVTAIEKTMGRSVISVILSLPLEESVYESLHE 477

RESULT 3
US-09-884-696-39
: Sequence 39, Application US/09884696
: Publication No. US20030035809A1
: GENERAL INFORMATION:
: APPLICANT: GEORGE, LISLE W
: APPLICANT: ANGELOS, JOHN A
: APPLICANT: HESS, JOHN F
: TITLE OF INVENTION: MORAXELLA BOVIS CYTOTOXIN, CYTOTOXIN GENE, ANTIBODIES
: TITLE OF INVENTION: AND VACCINES FOR PREVENTION AND TREATMENT OF MORAXELLA
: TITLE OF INVENTION: BOVIS INFECTIONS
: FILE REFERENCE: 481.06
: CURRENT APPLICATION NUMBER: US/09/884.696
: CURRENT FILING DATE: 2001-06-19
: NUMBER OF SEQ ID NOS: 41
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 39
: LENGTH: 478
: TYPE: PRT
: ORGANISM: Actinobacillus pleuropneumoniae
US-09-884-696-39

Query Match 12.1%; Score 262; DB 9; Length 478;
Best Local Similarity 24.2%; Pred. No. 2e-13;
Matches 111; Conservative 86; Mismatches 178; Indels 84; Gaps 18

QY 19 IGYGIIALTEVGLLGMMAFAPLDSAVIANGVVS-AEVSODVOHLEGMLAKILVREGKRY 77
Db 60 IAYLMLFLFLAV-ISTYSHVEIYATATGKALFSDRSKEIKFIENALVKELFVODGQRY 118
QY 78 KAGVYLFELDPTQANAAGITFRNOYVALKAMEAR--LLAE---RDQPSISFPAD--LTS 130
Db 119 EKDQGLLHHTALGNADADQOKTKSSLSTFKLERYRVEILLEVAADRLPLIELTKDEFKHA 178
QY 131 QRADPVARAIADEQAQFTER-----RQTIOGVLDLMAAORLOYOSEIEGIDROT 180
Db 179 TEEDTRIRYLLTEFOEFAMQOKYOKYKELALQRRAEKQTVLANIRK-----YEGISR-- 230
QY 181 QGLKQLOGEIEE-LIDRLKLDKGLVPRRLALLEAR---AGLSSSIGRLTADRSKA 235
Db 231 -----VENERIKLIDKLFNSKSTSHDVLQENRIIEAVNELAVYKSRINEYESDL 281
QY 236 VQASDPTOLKVNQIOEFEEQVQSITETRVRLAEVTEKEVVASDAQRIKIV--SPVNG 293
Db 282 RQAKEIHILITQLFRADILKIKQNV-EAKQDLSLEL-----NEQRDIASYIRAPVSG 335
QY 294 TAQNLRFETEGAVVRAAEPLVDIADBEDEAFVIAQHFOPTDVNVHMGVTEYRLPAFHS 353
Db 336 TVQQLKTHTVGGVVTTAETLWVIAPEDEVLLEVLTALIQNDIGFIEVGDAVIAKVFET 392
QY 354 GNPDERHDPA-----VADRTSDQQAARLEGLVARDVYKLRPHLGR--- 398
Db 393 -----PYTRGYLGMKYKNITLLEAHENQ-LGVFNSTIISIDRKT---LSGKQK 438
QY 399 ---VTAGMPACVIVPTGERTVLOYLFSPRLDRLRTTMR 434
Db 439 EIELSGMSVTAIEITGERSVISTYLLSPLEBSVESLRE 477

RESULT 4
US-09-884-696-38
: Sequence 38, Application US/09884696
: Publication No. US20030035809A1
: GENERAL INFORMATION:
: APPLICANT: GEORGE, LISLE W
: APPLICANT: ANGELOS, JOHN A
: APPLICANT: HESS, JOHN F
: TITLE OF INVENTION: MORAXELLA BOVIS CYTOTOXIN, CYTOTOXIN GENE, ANTIBODIES
: TITLE OF INVENTION: AND VACCINES FOR PREVENTION AND TREATMENT OF MORAXELLA
: TITLE OF INVENTION: BOVIS INFECTIONS

```

Query Match	6.6%;	Score 142;	DB 10;	Length 355;
Best Local Similarity	21.2%;	Pred. No. 0.00082;		

Query Match	6.0%;	Score 130;	DB 10;	Length 285;
Best Local Similarity	21.2%;	Pred. No. 0.0057;		

Matches 87; Conservative 56; Mismatches 103; Indels 164; Gaps 20;

```

23 IIALTFVGLGMAFA-----PLDSAVIANGVSAEVSODVOHLEBGMILVREGEKVK 78
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
14 VAAAVLAGIMRYMLSPWTRDARADVVY---VADYV---SGWTDLEVKONQYVK 66
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
79 AGCVLFELDPYQAAAAGITNQYVALKAMEARLLAEKDPSPISFPADLTQSRADPNVA 138
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
67 VGDVLMRID-----OERYQANLEO-----A 86
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
139 RAIDEAQOAFERRQTIQGVYDLMNAQRLQYSEIEGIDRQTQGLKQDGFTEDELIDLR 198
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
87 RAVAFETRQ-----QY-----LLRONEAR----- 106
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
199 KLYDKGLVPRRLALEARAGSLSGISGRLTADRSKAVOGASDTOLKVRQIKOEFEEGVCS 258
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
107 -----RSRL-----GIGASIAEDKE-----MNOINAAIRSEYQELA 139
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
259 QSITEVRRLAEVEKEVVASDAQRKIIVSPVNGTAONLR-----FTEGAVVRAAEPL 313
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
140 Q-----VKIAELN-----LKRSELRAARGOVNLELAQGNVATACQAVMA---L 181
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
314 VDIAPEDAEFYQAFOPPTDVNVMGVTEVRLPAFHSAGNPDPERHDPVA---VADRIS 371
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
182 VD-----QOSFYVAYFEETKLPGRVGMRAQVRL---MSGDPIDGTYESISSGITTDRNS 234
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
372 DPORQ-----ARLFLGI-VRVYDVKQLPPLHGRVTAQMPAOYIV 409
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
235 TPDQQLANVEPTFNWVRLAQRIPIVRIQLDQVPADV---HLSAGWTASTV 282
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 7
US-09-820-843A-113
Sequence 113, Application US/09820843A
Publication No. US2003039963A1

```

GENERAL INFORMATION:
APPLICANT: Council of Scientific and Industrial Research
TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PROTEIN
FILE REFERENCE: 063915
CURRENT APPLICATION NUMBER: US/09/820, 843A
CURRENT FILING DATE: 2001-03-30
NUMBER OF SEQ ID NOS: 118
SOFTWARE: PatentIn version 3.0
SEQ ID NO 113
LENGTH: 2354
TYPE: PRT
ORGANISM: L. major
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: AC005802_5 L6202.3
NAME/KEY: misc_feature
OTHER INFORMATION: g116899670
US-09-820-843A-113

```

Query Match 5.98; Score 128.5; DB 9; Length 2354;
Best Local Similarity 25.34; Pred. No. 0.15;
Matches 60; Conservative 31; Mismatches 97; Indels 49; Gaps 9;

```

101 QYVALKAMEARLLAEERDQ-RPSISFPADLTQSORAD-----PWVARIADE---QA 146
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
727 QRALEAGVAVRLAADRDEARQOLAAANEELQRLDTATQRAELEAQIARLAADRDEARQ 786
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
147 QFERRQTIQGVYDLMNAQRLQYSEIEGIDRQTQGLKQDGFTEDELIDLRKLYDKGLV 206
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
787 QLANAEELQRLDTATQRAELEAGVAVRLAADRDEARQOLAAANEEL----- 834
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
207 PRPRLALEARAGSLSGISGRLTADRSKAVOGASDTOLKVRQIKOEFEEVQSITETRV 266
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
835 -QORLDTATQORALEAVARLADRDARQ-----QIAA---NAEELQRLDTATQORA 885
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
267 RL-AEYVEKEVVASDAQKRIIVSPVNGTAONLRFFTEGAVVRAAEPLVDIAPDEA 322
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

DB 886 ELEAVARLAAANEELQRL-----DTATQORALEAVARLAA-----ADRDEA 929

```

RESULT 8
US-09-991-496-120
Sequence 120, Application US/09991496
Patent No. US20020169285A1
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Campos-Neto, Antonio
APPLICANT: Webb, John R.
APPLICANT: Dillon, Davin C.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Bhatia, Ajay
APPLICANT: Colier, Rhea
APPLICANT: Probst, Peter
APPLICANT: Brannon, Mark
TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
THERAPY AND DIAGNOSIS OF LEISHMANIASIS
FILE REFERENCE: 210121.420C9
CURRENT APPLICATION NUMBER: US/09/991,496
CURRENT FILING DATE: 2001-11-20
NUMBER OF SEQ ID NOS: 137
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 120
LENGTH: 2310
TYPE: PRT
ORGANISM: Leishmania major and chagasi
US-09-991-496-120

```

Query Match 5.98; Score 127; DB 9; Length 2310;
Best Local Similarity 23.48; Pred. No. 0.2;
Matches 75; Conservative 46; Mismatches 128; Indels 72; Gaps 13;

```

40 LDSAVIANGVSAEVSODVOHLEBGMILVREGEKVK-----AGVLFELDPYQAAA 94
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1318 LDTATQORALEAVARLAA-LEADGDEARQOLAAANEELQRLDTATQORALEAVARLA 1375
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
95 AG-----ITRNOYVALKAMEARLLAEERDQ-RPSISFPADLTQSORAD----- 134
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1376 ANAEELQRLDTATQORALEAVARLAAADRDEARQOLAAANEELQRLDTATQORALE 1435
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
135 PWVARIADE---QAFERRQTIQGVYDLMNAQRLQYSEIEGIDRQTQGLKQDGFIE 191
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1436 AQVARLADRDARQOLAAANEELQRLDTATQORALEAVARLADGDEARQOLANA 1495
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
192 DELIDLRKLYDKGLVPRRLALEARAGSLSGISGRLTADRSKAVO---GASDTOLKVRQI 249
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1496 EEL-----QORLDTATQORALEAVARLAAADRDEARQOLAAANEEL----- 1537
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
250 KOEFEEVQSITETRVRLAEVT-----EKEVVASDA---QRIKIVSPVNGTAONLRFF 301
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1538 -QORLDTATQORALEAVARLADGDEARQOLAAANEELQRL-----DTATQORAE 1589
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
302 TEGAVVRAAEPLVDIAPDEA 322
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1590 LEAVARLAA-----ADRDEA 1604
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 9
US-09-820-843A-114
Sequence 114, Application US/09820843A
Publication No. US2003039963A1
GENERAL INFORMATION:
APPLICANT: Council of Scientific and Industrial Research
TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PRO
FILE REFERENCE: 063915
CURRENT APPLICATION NUMBER: US/09/820, 843A
CURRENT FILING DATE: 2001-03-30
NUMBER OF SEQ ID NOS: 118
SOFTWARE: PatentIn version 3.0
SEQ ID NO 114

```


Db 218 VAOSDAAKAAEQADAAIEERLADFCGLNDRSLNDVIGLLDERESLASAESELAQAR 277
QY 199 KLYDKGLVPRRLALEARAGSLSSIGRLT----- 229
Db 278 AAGDLEAV-----AAAEKAVAGLEQSIASKTSWPSQDOTYLOSTALEAEARRYASTTE 332
QY 230 -----ADRSKAVOGASDQ-----LKVKOIKKEFFEQVSOSITE 263
Db 333 ALEIAERLYIDSLGKVDSELAAQRAVAEASHAQQDAAALGLETQOLSTOHOLEAQSSAID 392
QY 264 TRVRLAEYTEKEVVAQDAQKRIKI-----VSPVNGTAQNLFFETGAVVRAAEPLVDIAP 318
Db 393 AALGLASVDNKAATRSQGLKMDINNTVRSPTSVSSVO-AAQGO--PAAGALLSYA- 448
QY 319 EDEAFVIOAHFOPTDVNVHGMWTEVRLPAFHSAGNDP---PERHDPVAVA 367
Db 449 DDSEIKITANVKEAISNVTIGSRVTFPTP---STGTFEFAGRVSKVPIAAA 498

RESULT 12

US-09-815-242-5064
; Sequence 5064, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5064
; LENGTH: 2472
; TYPE: PRF
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-5064

Query Match 5.3%; Score 115.5; DB 10; Length 2472;
Best Local Similarity 20.6%; Pred. No. 1.9;
Matches 112; Conservative 78; Mismatches 178; Indels 175; Gaps 25;

QY 37 FAPLDSAVIA-----NGVSAEVSQDVHLEGG-MLAKI-----LYRE----- 73
Db 1534 FDIILDSAAALQRMDDVDVNTIELEALQDLHTLKGARMAIEIGLAELEFLEGL 1593
QY 74 -GEVVKAGQVLF-----ELDPTQANAAAGITRNQVVALKANAEARLLAERDQRPSTSF 124
Db 1594 CGGRLRASPALEFGLLOCHDELAEMLAEVARGHRTLPPGQALIAEIRRLRSDPDOLSYV- 1652
QY 125 PADLTSORADPMVARALADEQAO-----FTERRQ----- 153

Db 1633 ----TSVSLKPLAKGAADASELIDFLEADLLENLELALGHWGNGDQAPLDDL 1708
QY 154 -----TIOGVDMN-----AQR-----LOYOSKIEGIDROTQ 182
Db 1709 RILHTLKGARLARGTLEIGNLAHDECHLTDAQOGGAWPWSLLIDAGSGLEGLOQYD 1768
QY 183 LKDOIETIEDELIDKLYDKGLVPRRLALEA---RAGS-----LSGSGRLTA 230
Db 1769 LRERLA--EDD-----EAGEREPQALVQADDTDRAVASALAEITRLAPAGAIMA 1818
QY 231 DRSKAVOGASDQOLKVKROIKEEFFQVSOSTERFVR--AVTEKEV---VASDAQRI 285
Db 1819 AEAPPAAPATLTPVRAQEAQEAASRRAPQDELKVPABULELVNLAGETSTFGKRV 1878
QY 286 -KIVSPVNGTAQNLFFETGAVVRAAEPLVDIAPEDEAFVIOAH-----FOPTD 333
Db 1879 EQQVSDVQFTLGEH-----ESTIERVRODLRLDLETQOQLISBHQDAERAGYEEDFLE 1934
QY 334 VDNVHGMWTEVRLPAFHSAGNDPDERHDPVAVADRISDPQOKARFL- 381
Db 1935 MD--RYSOLOQLSRALFESAS-----DLIDLKETTAAKNRDAETLLLOQARVNTLQ 1985
QY 382 GIVR---VDVROLPPHLR---GRVTAGMPAOVIVPTG-----ERTVLOYLFSPLRDTLR 429
Db 1986 GLMRTMVPFDRLVPRRLRRIYQVAGELGQVEFVGNADGEMRTVLERIVAPLEHMLR 2045
QY 430 TTM 432
Db 2046 NAV 2048

RESULT 13

US-09-815-242-11146
; Sequence 11146, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11146
; LENGTH: 390
; TYPE: PRF
; ORGANISM: Haemophilus influenzae
US-09-815-242-11146

Query Match 5.3%; Score 114.5; DB 10; Length 390;
Best Local Similarity 20.4%; Pred. No. 0.16;

